

GENOME EDITING OF THE IL-7 RECEPTOR
GENE LOCUS USING TALENS

by
GÜLPERİ YALÇIN

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfillment of
the requirements for the degree of
Master of Science

Sabancı University
August 2014

GENOME EDITING OF THE IL-7 RECEPTOR
GENE LOCUS USING TALENS

APPROVED BY:

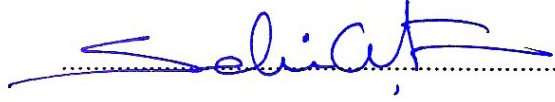
Assoc. Prof. Dr. Batu Erman
(Thesis Supervisor)



Prof. Dr. Canan Atılgan



Prof. Dr. Selim Çetiner



DATE OF APPROVAL:21.07.2014.....

© Gülperi Yalçın, 2014
ALL RIGHTS RESERVED

ABSTRACT

GENOME EDITING OF THE IL-7 RECEPTOR GENE LOCUS USING TALENS

Gülperi Yalçın

Biological Sciences and Bioengineering, MSc. Thesis, 2014

Thesis supervisor: Batu Erman

Keywords: IL-7R alpha, Glucocorticoid Receptor,
Transcription activator-like effector, TALEN, Genome Editing

IL-7 signaling is key to lymphocyte development and function in the mammalian immune system. In the first part of this study, we targeted the IL-7R alpha gene locus, encoding the IL-7 receptor protein, to identify its transcriptional control elements. We mutated two transcription factor binding sites in an evolutionarily conserved region containing a putative transcriptional enhancer by generating transcription activator like effector nucleases (TALENs) targeting these sites. We designed and constructed two pairs of TALENs targeting the glucocorticoid receptor (GR) and Notch binding sites in this region. We also targeted the exon 2 and exon 3 of IL-7R to delete a transcriptional control element in intron 2. We expressed these TALENs in the murine RLM11 (IL-7R positive) cell line and generated insertion and deletion mutations in the targeted sites. We used restriction fragment length polymorphism (RFLP) assays and DNA sequencing to detect the induced mutations and assessed their effects on IL-7R gene expression. We demonstrate that mutations induced in the GR transcription factor binding site do not reduce IL-7R gene expression, while mutations in the Notch binding site lower expression. In the second part of the study we directly targeted the gene encoding the GR transcription factor. We designed a TALEN pair targeting the translation start site to knockout gene expression. We introduced these TALENs into the human HCT116 cell line and performed RFLP assays to detect mutations. Our experiments demonstrate that TALENs can be used for genome editing to study gene transcription regulatory regions.

ÖZET

IL-7 ALMAÇ GENİNDE TALEN PROTEİNLERİ İLE GENOM MÜHENDİSLİĞİ

Gülperi Yalçın

Biyoloji Bilimleri ve Biyomühendislik, Master Tezi, 2014

Tez Danışmanı: Batu Erman

Anahtar Kelimeler: IL-7R alfa, Glukokortikoid Almacı,
Transcription activator-like effector, TALEN, Genom Mühendisliği

IL-7 sinyalleme sistemi memeli bağışıklık sisteminde lenfosit gelişimi ve fonksiyonu için anahtar konumundadır. Bu çalışmanın ilk bölümünde IL-7R alfa gen bölgesini hedefleyip gen ifadesini kontrol eden faktörleri tespit etmeyi amaçladık. Transcription activator-like effector nükleaz (TALEN) proteinleri ile enhancer bölgesindeki iki transkripsiyon faktör bağlanma bölgesinde mutasyonlar oluşturduk. Glukokortikoid almacı (GR) ve Notch bağlanma bölgelerini hedefleyen iki çift TALEN ile birlikte 2. ve 3. ekzon bölgelerini hedefleyen TALEN'ler oluşturup aralarındaki intron bölgesindeki transkripsiyon kontrol elementlerini genomdan silmeyi planladık. Bu TALEN'leri fare RLM11 (IL-7R+) hücre hattında ifade edip RFLP yöntemi ve DNA sekanslaması ile oluşturduğumuz mutasyonları tespit ettik. Mutant hücre hatlarının IL-7R ifadelerini analiz ederek Notch bağlanma bölgesindeki mutasyonların gen ifadesinde azalmaya yol açtığını, fakat GR bağlanma bölgesindeki mutasyonların bu ifade seviyesini düşüremediğini gözlemledik. Çalışmanın ikinci bölümünde doğrudan GR transkripsiyon faktörünün gen bölgesini hedefledik. Translasyon başlangıç bölgesinde çift sarmallı kesik oluşturacak bir TALEN çifti tasarlayıp gen “knockout” yapmayı amaçladık. Bu TALEN'leri insan HCT116 hücre hattında ifade edip RFLP ile mutasyonları tespit ettik. Deneylerimiz TALEN teknolojisi ile genomun modifiye edilip gen ifadesini kontrol eden faktörlerin çalışılabileceğini göstermektedir.

To my family...

ACKNOWLEDGEMENT

First of all, I would like to thank my supervisor Assoc. Prof. Dr. Batu Erman for his continuous guidance, support, and patience with me as I completed my master's degree. It has been a great pleasure to be a part of Erman Lab; I am very grateful for the provided opportunities and the unique project that I could work on during my two years of graduate education. I am forever indebted.

I also would like to thank my thesis committee, Prof. Dr. Canan Atılgan and Prof. Dr. Selim etiner for their advices and helpful criticism for my thesis evaluation.

My colleagues, my advisors and most importantly, my friends Canan Sayitoęlu, Bahar Shamloo, Nazlı Keskin, Emre Deniz, Őeyda Temiz and Ahsen zcan had been the reason I could complete this period with ease; by learning and having fun at the same time. I consider myself lucky for knowing them and for being able to work with them in our lab.

I cannot forget the importance of my comrades Esra nsal and Sıla zdemir in my life; even after the graduation they walked the same path as me and had been with me; the moral support they have given me is priceless.

I am always thankful to my parents and my lovely sisters, as their existence is the most precious treasure in my life. Just a few words would not be enough to explain in how many ways they provided for me; I could be the person I am only because they were there for me.

Finally, I would like to express my gratitude to the Scientific and Technological Research Council of Turkey, TBİTAK BİDEB for the support they gave me during my thesis project. Also, this project was supported by TBİTAK 109T315.

TABLE OF CONTENTS

1. INTRODUCTION.....	1
1.1. Transcription Activator like Effectors.....	1
1.1.1. Structural Features of TAL Effector Proteins.....	1
1.1.2. Crystal Structure of TAL Effector Proteins.....	4
1.1.3. Designing Custom TAL Effector Proteins.....	6
1.1.4. Targeted Genome Modification Using TALENs.....	7
1.1.5. Types of Genome Modification.....	10
1.1.6. Applications of Genome Editing Using TALENs.....	12
1.2. Interleukin-7 Signaling.....	13
1.2.1. Interleukin-7 and Interleukin-7 Receptor.....	13
1.2.2. IL-7 Receptor Signaling Pathways.....	14
1.2.3. Importance of the IL-7R Signaling for Lymphopoiesis.....	15
1.2.4. Regulation of the IL-7R alpha Gene.....	17
1.2.4.1. Notch Transcription Factor.....	18
1.2.4.2. Glucocorticoid Receptor (GR).....	20
2. AIM OF THE STUDY.....	22
3. MATERIALS AND METHODS.....	23
3.1. Materials.....	23
3.1.1. Chemicals.....	23
3.1.2. Equipment.....	23
3.1.3. Buffers and Solutions.....	23
3.1.4. Growth Media.....	24
3.1.4.1. Bacterial Growth Media.....	24
3.1.4.2. Mammalian Cell Culture Growth Media.....	24
3.1.5. Cell Types.....	25
3.1.6. Commercial Molecular Biology Kits.....	25
3.1.7. Enzymes.....	25
3.1.8. Vectors and Primers.....	26
3.1.9. DNA Molecular Weight Marker.....	29
3.1.10. DNA sequencing.....	29
3.1.11. Software and Computer Based Programs.....	29
3.2. Methods.....	30

3.2.1. Bacterial Cell Culture.....	30
3.2.1.1. Bacterial Cell Culture Growth.....	30
3.2.1.2. Competent cell preparation and transformation.....	30
3.2.1.3. Plasmid DNA isolation.....	31
3.2.2. Vector Construction.....	31
3.2.2.1. The General Methods Used in Vector Construction.....	31
3.2.2.2. Vector Construction for Homologous Recombination.....	32
3.2.3. Construction of TALEN expression Vectors.....	37
3.2.3.1. Identification of TALEN target sites.....	37
3.2.3.2. Assembly of custom TALEN constructs using Golden Gate TALEN Kit.....	38
3.2.4. Mammalian Cell Culture.....	44
3.2.4.1. Maintenance of Mammalian Cell Lines.....	44
3.2.4.2. Transient Transfection of Suspension Cells.....	44
3.2.4.3. Transient Transfection of Adherent Cells with PEI.....	45
3.2.4.4. Flow Cytometric Analysis.....	45
3.2.5. TALEN Induced Mutation Screening.....	46
3.2.5.1. Genomic DNA extraction.....	47
3.2.5.2. Restriction Fragment Length Polymorphism (RFLP) Analysis.....	47
3.2.5.3. Single Cell Analysis.....	48
4. RESULTS.....	50
4.1. Targeting IL-7R Gene.....	50
4.1.1. Targeting Transcription Factor Binding Sites of IL-7R Gene.....	50
4.1.1.1. TALENs Targeting GR Binding Site of IL-7R.....	51
4.1.1.1.1. Assembly of GR Site Targeting TALENs.....	52
4.1.1.1.2. Expression of the Constructed GR TALEN Pair in RLM11 cells and Detection of Site Specific Mutations.....	58
4.1.1.1.3. Single Cell Screening of GR TALEN transfected RLM11 cells and Detection of Mutants.....	60
4.1.1.1.4. IL-7R Expression in GR binding site Mutant RLM11 cells.....	61
4.1.1.2. TALENs Targeting Notch Binding Site of IL-7R.....	63

4.1.1.2.1.	Expression of a previously designed Notch TALEN pair in RLM11 cells and mutation screening via expression of IL-7R in single cell colonies.....	64
4.1.1.2.2.	Detection of mutation in suspected Notch TALEN transfected single cell colonies.....	66
4.1.2.	Use of TALENs to Delete an Entire Intronic Region of IL-7R.....	69
4.1.2.1.	Assembly of TALENs targeting Exon 2 and Exon 3 of IL7R gene.....	70
4.1.2.2.	Expression of Exon2 and Exon3 TALEN pairs in RLM11 cells and detection of mutation.....	75
4.1.2.3.	A map of the donor plasmid targeting the ECR1 region between Exon 2 and 3 by homologous recombination.....	77
4.1.2.4.	Simultaneous transfection of the Exon2, Exon3 TALENs and the donor dsDNA in RLM11 cells.....	78
4.2.	Targeting the GR gene.....	80
4.2.1.	TALENs Targeting human GR gene to induce Knock-In and Knock-Outs.....	80
4.2.2.	Assembly of TALENs targeting start site of hGR gene.....	81
4.2.3.	Expression of hGR TALEN pair in HCT116 cells and detection of mutation.....	85
4.2.4.	Construction of homologous donor plasmid to insert Venus gene into hGR endogenously.....	85
5.	DISCUSSION.....	91
6.	CONCLUSION.....	98
	REFERENCES.....	100
	APPENDIX.....	106
	APPENDIX A: Chemicals Used in the Study.....	106
	APPENDIX B: Equipment Used in the Study	108
	APPENDIX C: Molecular Weight Marker.....	110

LIST OF FIGURES

Figure 1.1 TALE structure and DNA recognition code.....	2
Figure 1.2 Some of the most frequently seen polymorphisms in TAL effectors from <i>Xanthomonas</i> spp	3
Figure 1.3 Crystal structure of the natural TAL Effector protein, PthXo1.....	5
Figure 1.4 TALE based custom proteins can be used to target DNA.....	7
Figure 1.5 TALEN structure for genome editing.....	9
Figure 1.6 TALEN induced genome editing.....	11
Figure 1.7 The IL-7 receptor signaling pathway.....	15
Figure 1.8 IL-7R expression by lymphocytes.....	16
Figure 1.9 IL-7R gene locus with different transcription factor binding sites.....	17
Figure 1.10 Notch signaling.....	19
Figure 1.12 Glucocorticoid receptor signaling.....	21
Figure 3.1 The strategy for fusion of exon 2 and exon 3 of IL7R gene.....	33
Figure 3.2 The strategy for construction of puromycin resistance and Venus-YFP inserted hGR gene homologous plasmid.....	34
Figure 3.3 Gibson Assembly working principle for PCR products with homologous ends.....	37
Figure 3.4 Golden Gate assembly of custom TALE and TALEN constructs.....	39
Figure 3.5 Timeline for TALEN construction using TALEN Golden Gate kit.....	39
Figure 3.6 General strategy for detection of TALEN induced mutation at the target site.....	46
Figure 3.7 Methods for obtaining single cell colonies.....	48
Figure 4.1 Schematic representation of the mouse IL7R α gene locus.....	51
Figure 4.2 Binding sites for GR binding site targeting TALENs.....	52
Figure 4.3 Plasmid maps showing GR TALEN2 Left and Right TALEN pair after Golden Gate Reaction #1.....	53
Figure 4.4 Plasmid maps showing GR TALEN3 Left and Right TALEN pair after Golden Gate Reaction #1.....	54

Figure 4.5 Agarose gel images of colony PCR and control digests of GR TALEN2 and TALEN3 pairs after Golden Gate Reaction #1.....	55
Figure 4.6 Plasmid maps showing fully assembled GR TALEN2 Left and Right TALEN pair in pC-GoldyTALEN backbone.....	56
Figure 4.7 Plasmid maps showing fully assembled GR TALEN3 Left and Right TALEN pair in pC-GoldyTALEN backbone.....	57
Figure 4.8 Agarose gel images showing colony PCR and control digest results of GR TALEN2 and GR TALEN3 pairs after Golden Gate reaction #2.....	58
Figure 4.9 IL-7R expression levels of GFP and GR TALEN3 transfected RLM11 cells in compared to untransfected cells.....	59
Figure 4.10 RFLP analysis on the GR TALEN transfected and untransfected RLM11 cells.....	59
Figure 4.11 RFLP results for GR TALEN transfected single cell colonies.....	60
Figure 4.12 Selected RLM11 single cell colonies to send sequencing.....	60
Figure 4.13 Sequencing results of single cell colonies that had uncut bands in the RFLP assay.....	61
Figure 4.14 FACS analysis showing IL7R expressions of mutant RLM11 single cell colonies.....	62
Figure 4.15 The binding sites for IL-7R Notch site targeting TALEN pair.....	63
Figure 4.16 IL-7R expression levels of Notch TALEN transfected RLM11 cells in compared to untransfected cells.....	64
Figure 4.17 RFLP analysis on the Notch TALEN transfected and untransfected RLM11 cells.....	64
Figure 4.18 IL-7R expression levels of Notch TALEN transfected RLM11 single cell colonies that have lower expression levels in compared to WT.....	65
Figure 4.19 RFLP assay for Notch TALEN transfected RLM11 single cell colonies.....	66
Figure 4.20 IL-7R expression levels for Notch site mutated RLM11 cells.....	66
Figure 4.21 The sequencing results for Notch site mutated single cell colonies.....	67
Figure 4.22 IL7R gene ECR2-ECR3 locus.....	68
Figure 4.23 PCR amplification of 74 th colony with different primers.....	68
Figure 4.24 The strategy to delete ECR1 region of the IL-7R gene using two TALEN pairs and a homologous donor DNA.....	69
Figure 4.25 Binding sites of TALENs targeting Exon 2 of IL-7R gene.....	70

Figure 4.26 Binding sites of TALENs targeting Exon 3 of IL-7R gene.....	71
Figure 4.27 Plasmids maps of Exon 2 targeting TALEN pair in pFUS_A and pFUS_B intermediary plasmids.....	71
Figure 4.28 Plasmids maps of Exon 3 targeting TALEN pair in pFUS_A and pFUS_B intermediary plasmids.....	72
Figure 4.29 Agarose gel images showing colony PCR and control digest results of Exon 2 and Exon 3 targeting TALENs.....	73
Figure 4.30 Plasmid maps for fully assembled Exon 2 targeting TALEN pair in their final expression vector.....	73
Figure 4.31 Plasmid maps for fully assembled Exon 3 targeting TALEN pair in their final expression vector.....	74
Figure 4.32 Agarose gel images of colony PCR and control digest results of both Exon 2 and Exon 3 targeting TALENs.....	75
Figure 4.33 FACS analysis of Exon 2 and Exon 3 TALEN transfected RLM11 cells.....	76
Figure 4.34 RFLP assay of the Exon 2 and Exon 3 TALEN transfected RLM11 cells.....	76
Figure 4.35 The plasmid map for E2-E3 fusion product and the agarose gel image of the control digest.....	77
Figure 4.36 RFLP assay for Exon 2 and Exon 3 TALEN pair co-transfection.....	78
Figure 4.37 FACS analysis of Exon 2 and Exon 3 TALEN transfection and cotransfection in RLM11 cells.....	79
Figure 4.38 The strategy to knock out human glucocorticoid receptor using TALENs and insert Venus and puromycin resistance genes using homologous recombination.	81
Figure 4.39 The binding sites of the hGR TALEN pair on the hGR gene start region.....	82
Figure 4.40 The plasmid maps for hGR right and left TALENs in pFUS_A and pFUS_B intermediary plasmids.....	82
Figure 4.41 Agarose gel images of the colony PCR results of hGR TALENs after Golden Gate reaction #1.....	83
Figure 4.42: Plasmid maps of hGR left and right TALEN pair in pC-GoldyTALEN backbone.....	84

Figure 4.43: Colony PCR and double digest control for hGR TALEN pair after Golden Gate reaction #2.....	84
Figure 4.44 RFLP assay for hGR TALEN transfected HCT116 WT cells.....	85
Figure 4.45 Left homologous arm PCR site on hGR gene locus.....	86
Figure 4.46 Right homologous arm PCR site on hGR gene locus.....	87
Figure 4.47 The optimized PCR results of left and right homologous arms of hGR gene. The regions were amplified from HCT116 genomic DNA extract.....	87
Figure 4.48 The plasmid map and optimized PCR result for puromycin resistance gene.....	88
Figure 4.49 The plasmid map and optimized PCR result for Venus gene.....	89
Figure 4.50 The experiments done to fuse PCR products of the hGR donor plasmid.....	90
Figure 4.51 The plasmid map for the hGR donor construct.....	90

LIST OF TABLES

Table 3.1 List of vectors used in this project.....	26
Table 3.2 List of primers used in this project.....	27
Table 3.3 List of software and computer based programs used in this study.....	29
Table 3.4 The components and the amounts they were used for separate PCR reactions that were done in hGR donor plasmid construction.....	35
Table 3.5 Optimized PCR conditions for Venus-Right Arm fusion.....	35
Table 3.6 Optimized PCR conditions for Left Arm - Puromycin resistance gene fusion.....	36
Table 3.7 Binding sites of TALEN pair and spacer sequences on the 5'-3' coding strand.....	38
Table 3.8 Components and amounts for Golden Gate reaction #1.....	40
Table 3.9 Optimized colony PCR conditions.....	41
Table 3.10 Components for the first part of Golden Gate reaction #2.....	42
Table 3.11 Components for second part of Golden Gate reaction #2.....	42
Table 3.12 Optimized PCR conditions for TALEN target site amplification.....	47

LIST OF ABBREVIATIONS

α	Alpha
β	Beta
γ	Gamma
K	Kappa
CLP	Common lymphoid progenitor
CSL	CBF1/ RBPjk/ Su(H)/ Lag1
DN	Double Negative
DP	Double Positive
DSB	Double stranded break
ECR	Evolutionarily Conserved Region
ETP	Early Thymic Precursors
GABP	GGAA binding protein
GR	Glucocorticoid Receptor
GRE	Glucocorticoid Response Element
HR	Homologous Recombination
Hrp	Hypersensitive Response and Pathogenicity
HSC	Hematopoietic stem cell
Hsp90	Heat Shock Protein 90
IL	Interleukin
IL-7R	Interleukin-7 Receptor
INDEL	Insertion and Deletion
JAK	Janus Kinase
MAML	Mastermind-like

NICD	Notch Intracellular Domain
NF- κ B	Nuclear factor - kappa light chain enhancer of activated B cells
nGRE	Negative Glucocorticoid Response Element
NHEJ	Non-homologous End Joining
NK cell	Natural Killer Cell
NLS	Nuclear Localization Signal
P2A	Porcine Teschovirus-1 2A
PEI	Polyethylenimine
Runx1	Runt-related transcription factor 1
RFLP	Restriction Fragment Length Polymorphism
RVD	Repeat Variable Di-residue
SCID	Severe Combined Immunodeficiency
SP	Single Positive
STAT	Signal transducer and activator of transcription
T3S	Type III Secretion System
TAL	Transcription Activator-like
TALE	Transcription Activator-like Effector
TALEN	Transcription Activator-like Effector Nuclease
TBE	Tris-borate-EDTA
TCR	T-cell Receptor
TD	Translocation Domain
TLR-4	Toll-like Receptor 4
TNF- α	Tumor Necrosis Factor Alpha
YFP	Yellow Fluorescent Protein
ZFN	Zinc Finger Nuclease

1. INTRODUCTION

1.1 Transcription Activator like Effectors

Gram-negative bacterial plant pathogens of the *Xanthomonas* genus mostly owe their pathogenicity to the Hrp-Type III secretion (T3S) system which translocates effector proteins into plant cells. The largest effector family consist of transcription activator like (TAL) effectors which functions as transcription activators of plant genes. With the help of a nuclear localization signal TALE proteins translocate to the nucleus and bind targeted promoters in the host genome to perform a variety of tasks to promote bacterial infection, proliferation and dissemination ¹.

1.1.1 Structural Features of TAL Effector Proteins

A typical TALE structure is composed of an N-terminal translocation domain (TD), a central DNA binding domain (DBD), two nuclear localization signals (NLS) and a transcriptional activation domain (AD) in the C-terminal region. The DNA binding domain is made up of several tandem repeats and a final half-repeat as it is shown in Figure 1.1. Each repeat module generally is 34 amino acids (aa) in length; the last C-terminal half-repeat consists of 20 aa. The repeats are highly conserved; but the residues at positions 12 and 13 are polymorphic and called ‘repeat-variable di-residue’ (RVD). These two residues are responsible for recognizing and binding the specific DNA base pairs.

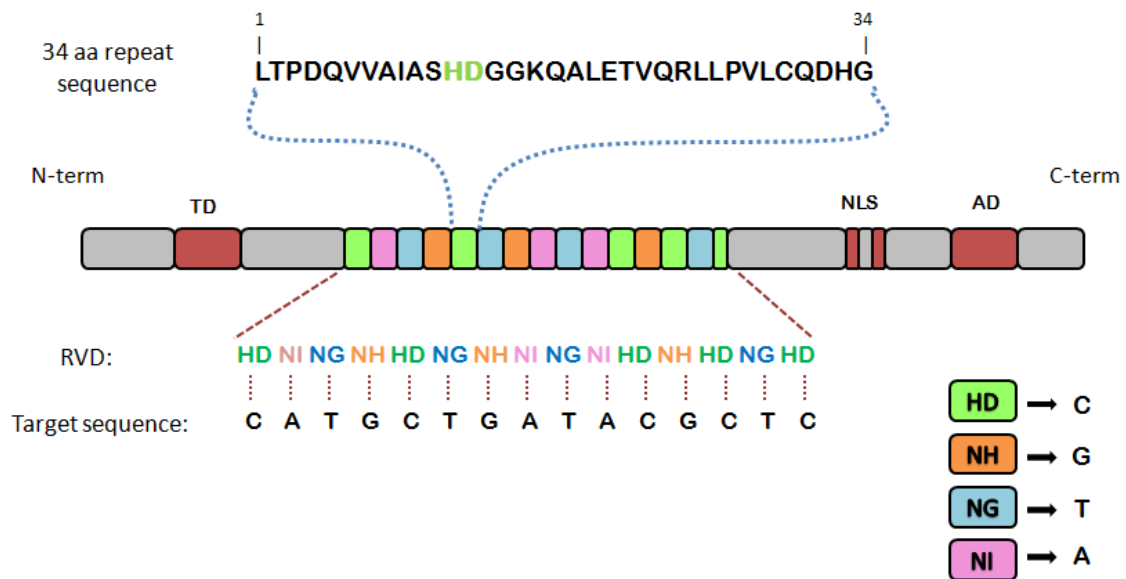


Figure 1.1 TALE structure and DNA recognition code. The transcription activator-like effector structure has an N-terminal translocation domain (TD), an array of repeats as a DBD, two NLS and an activation domain at the C-terminal. Each repeat in the DBD consists of the same 34 amino acid sequence with the exception of the 12th and 13th aminoacids, called repeat-variable di-residues (RVDs). Each RVD recognizes a different base of nucleotides. In the sketch, color coded four common RVDs and the target sequence they would bind is shown as an example ².

The number of repeats and their order changes the sequence TAL effectors bind, and this is mediated by a code. This code was deciphered by independent groups, and commonly found RVDs HD, NG/HG and NI has been proven to bind specifically to cytosine, thymine and adenine respectively. There are also other RVDs that are not exclusive, NN was found to recognize both adenine and guanine while NS could bind all four of the bases. Bioinformatic studies done with a large variety of TAL repeats has shown that the order of the repeats do not have a significant pattern and the binding specificity of one repeat does not affect the neighboring one ^{3,4}. However, most of the natural target sites had a conserved thymine in position zero (the base preceding the first base pair recognized by the central DBD) and the presence of T₀ was found to be necessary for full gene activation ⁴. Even though the region at the N-terminal of the first repeat did not seem to be conserved, later on a study done with the protein sequence of this region showed that the secondary structure is conserved to some degree and can specifically bind T ⁵.

Even though the predominant length of TAL repeats is 34 amino acids, in naturally existing TAL effectors, repeats of different sizes exist, and apart from the RVD there are polymorphisms observed in different residue positions. In *Xanthomonas* spp. polymorphisms at positions 4 and 32 are common (Figure 1.2), and in addition to these there are other amino acid substitutions which are rarely found. The polymorphisms in non-RVD domains do not seem to affect the base preference of the repeat and their function is not clear. Also, it is not known whether all types of repeats of various lengths and different polymorphisms that are found in nature are functional ¹.

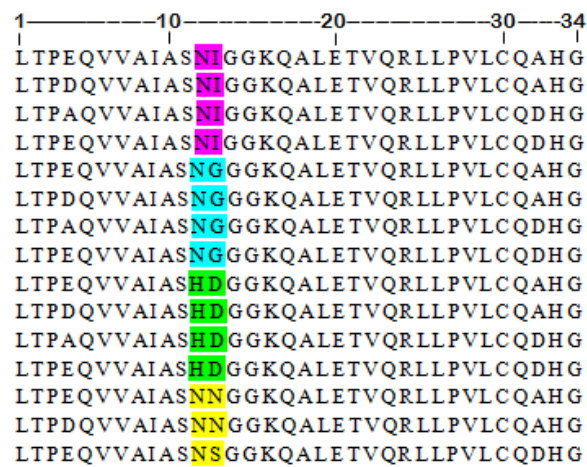


Figure 1.2 Some of the most frequently seen polymorphisms among 2023 TAL repeats of 113 known TAL effectors from *Xanthomonas* spp. Most TAL repeats consist of 34aa, the repeat units are generally conserved with the exception of 12th and 13th (RVD) along with 4th and 32nd positions, which do not seem to have an effect on base preference ¹. The base pairs these repeats recognize are shown with the colors indicated in Figure 1.1.

The simplicity of the one repeat-one base system and the independence of these repeats from each other in terms of specificity enables targeting any sequence in the genome by designing custom arrays of repeats and it enables the construction of artificial transcription factors as well as proteins with various functions by fusing a TAL protein to other domains ².

1.1.2 Crystal Structure of TAL Effector Proteins

The crystallization studies done on TAL effectors unraveled many features of TALE proteins at the atomic scale. In 2012, Mak et al. crystallized the DNA binding domain of the naturally occurring TAL effector protein PthXo1 from the rice pathogen *X. oryzae* as bound to its DNA target (PDB:3UGM) ⁵. In the same year Deng et al. shown the crystal structure of an artificially engineered TAL effector protein, dHax3 in both DNA-free and DNA-bound forms (PDB: 3V6P and 3V6T, for DNA-free and DNA-bound structures, respectively) ⁶. These two studies shed light on many questions about the sequence-specific recognition of DNA by the TAL repeat structure.

Crystallographic studies of TALEs showed that the DNA binding domain forms a right-handed superhelical assembly wrapped around the B-form DNA helix in a way which enables specific repeat-nucleotide interactions (Figure 1.3). An array of TAL repeats complete a full helical turn around the DNA, the RVD loops form the inner most spiral with a pitch of 60 Å per turn. The consecutive bases of the target site are intimately engaged in the major groove and all of the repeats in the DNA-bound structure form nearly identical two-helix bundles, reflecting the degree of high sequence conservation. The consecutive helices are packed left handedly and the shorter one, helix a, spans positions 3 to 11 while the longer and bent one, helix b is made up of residues 14 to 33. The two helices are connected by a short loop which contains the RVD. The kink in the second helix is generated by a proline at position 27 and it appears to be critical for the sequential packaging and association of tandem repeats with the DNA double helix ^{5,6}.

In both of the studies, the two hypervariable residues in the RVD loops are positioned in close proximity to the sense strand in the DNA major groove and the first and the second residue appear to have different biochemical roles. While the sequence-specific contacts are made exclusively by the 13th residue in each RVD, the 12th residue does not directly contact DNA. The first residue in the RVD makes direct hydrogen bonds to the backbone carbonyl oxygen of the alanine at position 8 in each repeat through its side chain, constraining the RVD containing loop. Ala8 is invariant and located at the C-terminal end of helix a in each TAL repeat, and appears to have a

critical role in stabilizing the conformation of the RVD loop by binding the 12th residue, thus facilitating the sequence recognition for the 13th residue ^{5,6}.

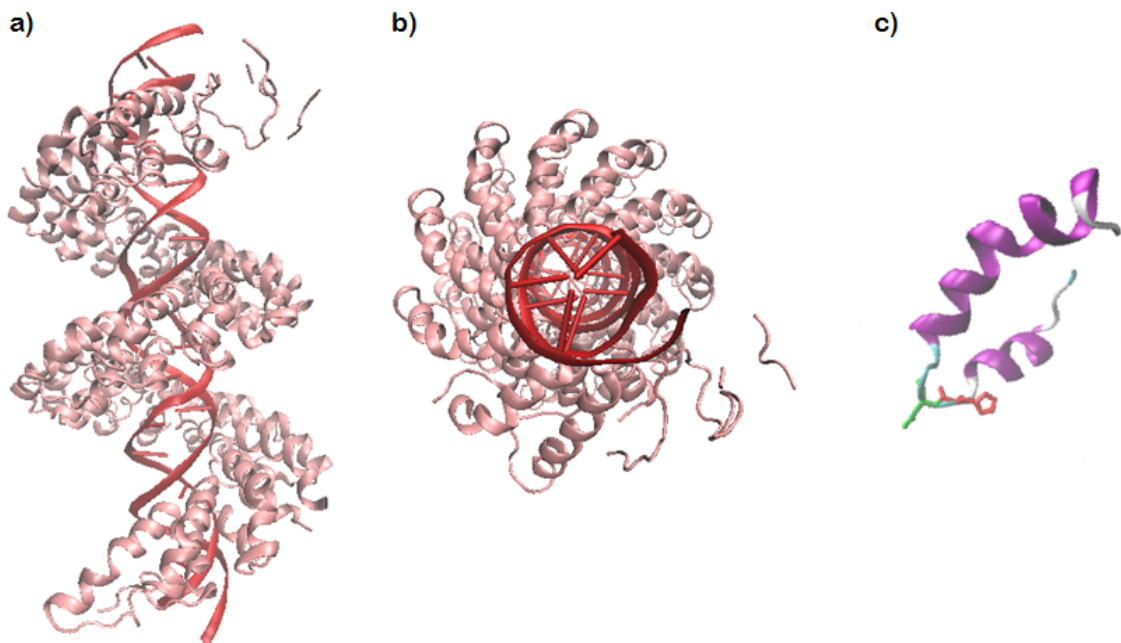


Figure 1.3 Crystal structure of the natural TAL Effector protein, PthXo1. a) Side view of PthXo1 in its DNA bound form. The protein backbone is indicated in pink and the DNA double helix is shown in red. b) Top view of the DNA bound PthXo1 c) Crystal structure of a single repeat unit containing an HD RVD, the H residue is shown in red, the D residue in green and alpha helices in purple (PDB: 3UGM ⁵).

All RVDs appear to have different characteristics to bind their corresponding bases, but their specificity is determined by the 13th residue. In the HD RVDs, the aspartate residue makes van der Waals contacts with the edge of the cytosine base and a hydrogen bond to the cytosine N4 atom. The contacts between the cytosine base and the charged acidic side chain of Asp13 are both physically and electrostatically specific and close out the possibility of binding other bases. In the case of NS, which is a non-selective RVD that binds all four of the bases, the hydroxyl group of Ser13 donates a hydrogen bond to the N7 atom of adenine, which also exists in guanine. Binding of cytosine and thymine might require a slightly different conformation of the loop. In the case of the NN RVD too, the second asparagine residue is positioned to make a hydrogen bond with the N7 nitrogen of the corresponding guanine base. Since the N7 nitrogen is available in both purines, NN can also bind adenine. NI RVD which binds an adenine base rather specifically demonstrates an exceptional contact pattern. The aliphatic side chain of the isoleucine residue was observed to make nonpolar van der

Waals contacts to C8 of the adenine purine ring which would normally cause desolvation of at least one polar atom in the ring. This might be the reason behind its reduced affinity. The contact made between NG and HG repeats and the thymine base is also interesting. Instead of making any specific interactions, the placement of Gly at position 13 allows sufficient space to contain the 5' methyl group of thymine. The distance between the C of Gly13 and the 5-methyl group of thymine is small (around 3.4 Å), allowing van der Waals interactions. Substitution of Gly with any other residue would likely introduce a steric clash with the 5-methyl group of thymine, explaining the specificity of NG and HG binding⁵⁻⁷.

The PthXo1 structure also revealed two degenerate repeat folds that are at the N-terminal of the central repeats that appear to cooperate to specify the conserved thymine that precedes the targeted sequence by RVDs. Residues 221 to 239 and residues 256 to 273 in the PthXo1 structure each forms a helix and an adjoining loop that resembles helix a and the RVD loop in the repeats of the DNA binding domain. These two N-terminal regions approach each other near the 5' thymine base, making a van der Waals contact with the methyl group of that base⁵.

1.1.3 Designing Custom TAL Effector Proteins

Thanks to the simplicity of the DNA recognition code and the modularity of the protein, TALEs allow the design of many functional proteins that would modify gene sequences and gene expression targeting any site in the genome. In principle, only by determining the number and order of the required RVDs, any site can be targeted; and by fusing these repeats to different functional domains, TALE proteins can be directed to their specific loci to induce the desired modification (Figure 1.4). Fusion of regulatory domains to TALE repeats enabled construction of proteins that induce activation or repression on the endogenous expression of the targeted genes^{8,9}. For site directed mutagenesis nonspecific nucleases fused to the TALE DNA binding domains have been used¹⁰⁻¹². Also, TALEs fused with the catalytic domains of invertase, named TALE recombinases, have become another important tool for site-specific modification of the genome¹³.

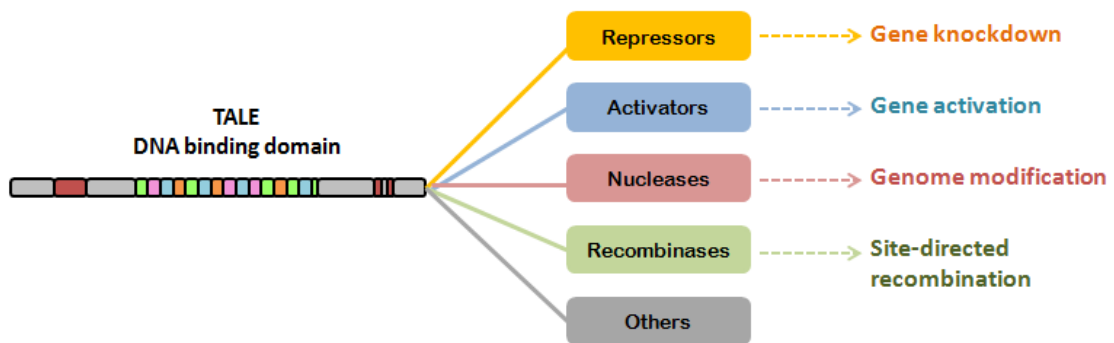


Figure 1.4 TALE based custom proteins can be used to target DNA. Functional domains such as activators, repressors, nucleases and recombinases can be fused to the central DNA binding domain of TAL effectors on the C-terminal end for targeted modification of genomes. TALE central repeats are color coded according to RVDs as shown in Figure 1.1 ².

Assembly of the TALE repeats are generally done with the most common RVDs NI, NG, HD and NN which bind adenine, thymine, cytosine and guanine respectively. Among these RVDs, NN also recognizes adenine which might be an impediment if specific recognition of a target region is desired, because it is possible for the designed protein to bind other sites in the genome. For that reason, NK and NH, RVDs that are found less common in nature, but more specific in recognizing guanine could be used; however, comparison studies have shown that even though these two were more specific, the TALE proteins that include these RVDs for guanine recognition had significantly lower activities compared to their NN containing counterparts ^{14,15}. Therefore, while constructing TALE proteins against targeted loci, both specificities and binding efficiencies of the individual RVDs should be taken into account and they should be chosen according to the requirements of the experiment.

1.1.4 Targeted Genome Modification Using TALENs

TALE nucleases (TALENs) are generated by the fusion of the DNA binding domain of a TALE protein to a nuclease functional domain. Using site specific nucleases for genome editing has become a trend in both studies of gene function and gene therapy research. Double stranded breaks induced by nucleases can result in gene inactivation because of non-specific DNA repair, and desired genome modifications can be induced by homology directed repair in specific sites. Before TALENs, zinc finger

nucleases (ZFNs) were being used as programmable and sequence specific tools for targeting endogenous gene loci and were utilized for therapeutic purposes that had proved to be a great success¹⁶. ZFNs were used to directly correct the disease causing mutations associated with X-linked severe combined immune deficiency (SCID) using homology directed repair¹⁷, to genetically repair Parkinson's disease-associated mutations within the SNCA gene in patient-derived human iPS cells¹⁸; and to knockout the CCR5 (C-C chemokine receptor type 5) gene in primary T cells by ZFN induced non-homologous end joining (NHEJ) to render these cells resistant to HIV. Since CCR5 is a co-receptor that takes part in HIV infection, this method holds great potential to defeat HIV and currently is being examined under clinical trials¹⁹.

TALENs are second generation genome editing tools. The first genome editing tool to be used was ZFN. Among DNA binding motifs, the zinc finger domain is the most common type found in eukaryotes. An individual zinc finger consists of approximately 30 amino acids in a conserved $\beta\beta\alpha$ configuration and several amino acids on the surface of the α -helix typically contact 3bp in the major groove of DNA with different selectivity levels. After the discovery of a highly conserved linker sequence, synthetic arrays that contain more than three zinc finger domains were developed and eventually custom zinc-finger proteins that recognize 9-18bp long DNA sequences were constructed²⁰. Generation of a zinc finger module library that contains the domains which recognize nearly all of the 64 possible nucleotide triplets led to the modular assembly of DNA binding domains that targets unique DNA sequences with relatively high specificity. The fusion of this domain to a DNA nuclease made this protein an important tool for targeted gene modification. Traditionally targeted gene inactivation, replacement, or gene insertion could only be achieved by homologous recombination which had too low efficiency in mammalian cells without the induction of double strand breaks; however, previously found site-specific nucleases would target multiple sites in complex genomes and their cytotoxicity levels were high²¹. For that reason, the zinc-finger motif was fused to the non-sequence-specific DNA cleavage domain of the restriction enzyme FokI, which is a type IIS restriction endonuclease that functions as a dimer. To induce double strand breaks, ZFNs that target two closely oriented inverted half sites should bind each other from their FokI domains at the spacer region in between the half sites. Therefore, DNA cleavage is generated only upon the heterodimerization of the two nucleases. Using this strategy for targeted genome

terminal end of the proteins and as the DBD of the second monomer binds the reverse strand of the DNA leaving a short distance in between, the FokI domains of the two proteins assemble on the spacer region and generate a double strand break (Figure 1.5)

11

TALEN technology was successfully used for targeted genome editing in yeasts²⁴, *C. elegans*²⁵, plants as model organisms or crops^{26,27}, *Drosophila melanogaster*²⁸, zebrafish²⁹, frogs³⁰, mouse^{31,32}, rat³³, and livestock³⁴. With the experiments done so far, the TALEN efficiency varies usually from 10 to >50% with an average around 22% cells mutated^{35,36}. In studies that are done to compare TALENs with ZFNs, along with the simplicity of its design and assembly, TALENs were found to be significantly more mutagenic and efficient³⁵ while the toxicity caused by off-target cleavage of the genome was much less frequent³⁷, rendering TALENs more preferable over ZFNs.

1.1.5 Types of Genome Modification

The basis of targeted genome modification through site-specific nucleases is to induce double strand breaks at the targeted site and to trigger cellular repair mechanisms. The double stranded break is either repaired by non-homologous end joining (NHEJ) or by homologous recombination (HR) in case a homologous donor DNA is supplied (Figure 1.6). After TALEN pairs induce a double strand break on the targeted site, NHEJ-mediated repair leads to the introduction of small insertions or deletions (INDELs) on the targeted site resulting in the disruption of the region, two-thirds of which causing frame-shift mutations that would knock-out the gene. TALENs were successfully used to knock out various genes without the introduction of exogenous DNA³⁴. Apart from random mutations, homologous recombination can result in gene deletions, gene insertions or gene replacements can be induced at the target site by co-delivering TALEN pairs with a donor plasmid that has locus-specific homologous arms along with the desired sequence. With this method, specific genes can be integrated into the genome or an epitope tag can be inserted to label the protein of interest, a defective gene can be repaired or a large sequence can be removed from the genome. Along with plasmids, linear DNA sequences and single-stranded oligonucleotides can be used as effective donors¹⁶. It was also shown that by the transfection of multiple TALEN pairs

simultaneously, deletion or inversion of large chromosomal segments were possible. In cases where different chromosomes were targeted, translocations could be induced³⁸. In addition, by synchronizing TALEN mediated DNA cleavage of the donor DNA with the chromosome, large expression cassettes (~15kb) could be inserted into the genome through NHEJ-mediated ligation³⁹.

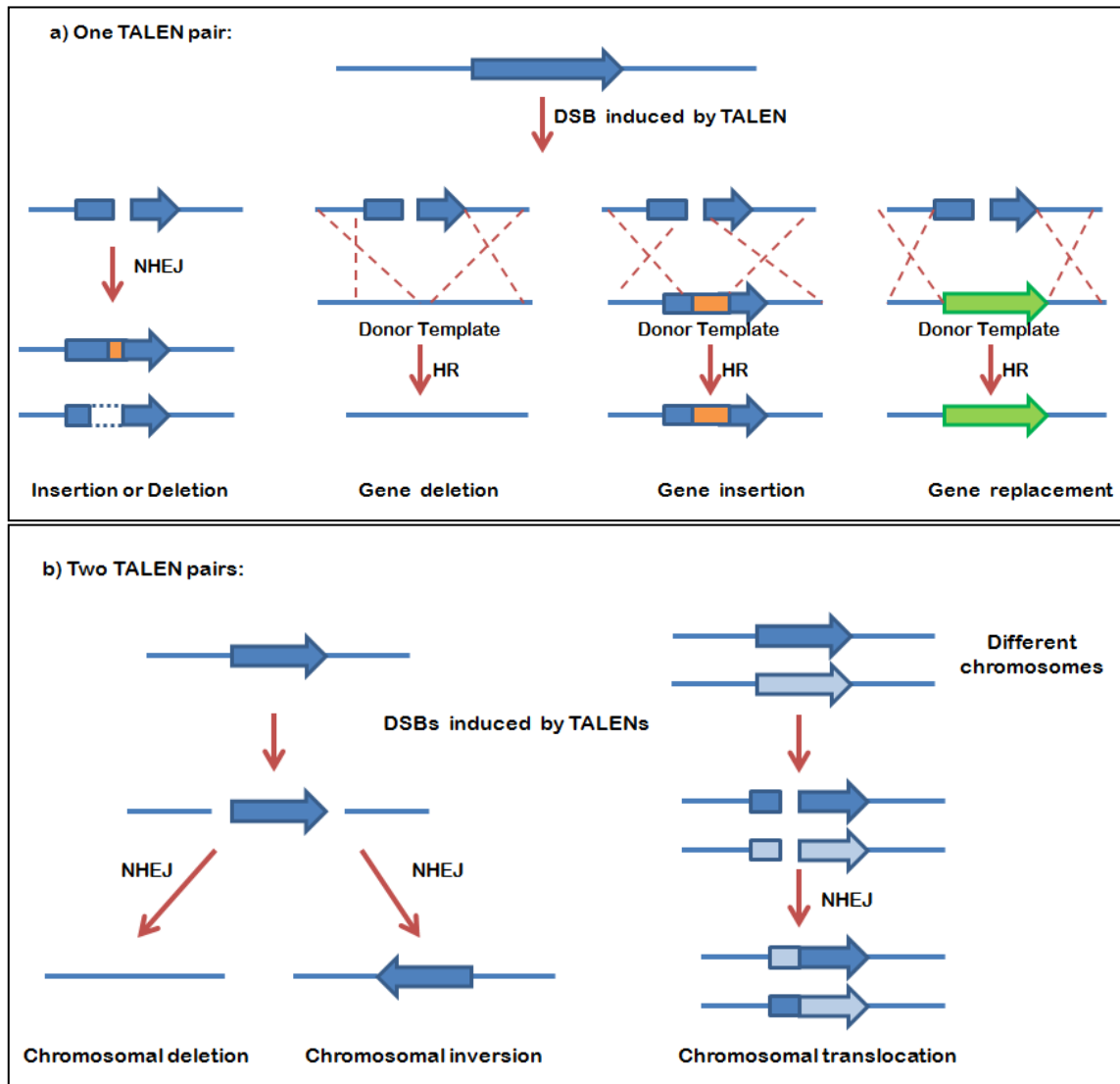


Figure 1.6 TALEN induced genome editing. Genome editing after DSB generation is done either by non-homologous end joining (NHEJ) or by homologous recombination (HR). a) In the case genome editing using one TALEN pair, NHEJ results in small insertions and deletions (INDELs) at the site of the DSBs. HR can be used for gene deletion, gene insertion (for example an epitope tag) or gene replacement (for example a fluorescent reporter gene such as GFP) depending on the donor template used. b) If two TALEN pairs create DSBs on the same chromosome simultaneously, NHEJ mediated repair may result in chromosomal deletion or inversion. If DSBs are generated on different chromosomes, translocations may occur^{2,23}.

1.1.6 Applications of Genome Editing Using TALENs

Targeted genome modification by TALENs have been used in various model organisms. First studies typically used NHEJ-mediated mutagenesis to knock out genes or to nullify a phenotype^{35,36}. TALEN pairs along with double strand donor DNA were used to insert expression cassettes into targeted regions in human cells⁴⁰ or induce specific modifications in the targeted site of zebrafish embryos by homologous recombination⁴¹. Similarly, homology directed repair could successfully be induced with single strand oligonucleotide donors that have ~50bp long homology arms leading to precise modifications in zebrafish and mouse models^{32,42}. This method can be used for fusion of endogenous genes to sequences encoding epitope tags or fluorescent reporter proteins such as GFP to track protein expression, distribution and interaction with other proteins (Figure 1.6)³⁸. Simultaneous transfection of two TALEN pairs generated heritable large chromosomal deletions in silkworm⁴³ and livestock genomes³⁴. All of these studies have shown that the TALENs are applicable for many means of gene editing. One of the most important approaches of genome modification studies is to develop knock-out model organisms and a widely used and successful way to do it is the microinjection of *in vitro* synthesized mRNAs encoding a custom TALEN pair into a zygote. With this method, zebrafish, livestock, rat and mouse disease models were generated^{32-34,41,42,44}.

In this study we used TALENs to disrupt two of the transcription factor binding sites in the enhancer region of the IL7R gene using NHEJ-mediated INDEL mutagenesis. In addition, we introduced two TALEN pairs simultaneously to delete an entire intronic region in the same gene. We also used TALENs against the translation start site of the GR gene to knock out gene expression and we designed a donor plasmid homologous to the same site with a Venus-YFP insert to fuse it with GR.

1.2 Interleukin-7 signaling

1.2.1 Interleukin-7 and Interleukin-7 Receptor

Interleukin-7 (IL-7) is a cytokine which plays an essential non-redundant role in the development, differentiation and survival of lymphocytes. IL-7 was first discovered in 1988 as a factor that promoted the growth of the murine B cell precursors in a bone marrow culture system⁴⁵. Then it was shown that injecting mice with IL-7 increased both T and B lymphocyte numbers dramatically while the studies with IL-7 and IL-7 receptor (IL-7R) defective models demonstrated significant reductions in the number of lymphocytes⁴⁶. These studies and the following ones confirmed IL-7's role in lymphocyte development and proliferation. In addition to these, IL-7 has a role in homeostasis of T lymphocytes; as well as in the early and late stages of the T cell development, it promotes cell survival in naïve and memory T cells of the peripheral immune system⁴⁷.

The human IL-7 gene is 72kb long and it is located on chromosome 8, encoding a protein of 20kD; whereas murine IL-7 gene is 41 kb long, located on chromosome 3 and encodes a protein of 18kD. Due to post-translational glycosylation, the active form of human IL-7 is 25kD in size and it is a single chain protein consisting of four α helices with a hydrophobic core. IL-7 is produced by non-lymphoid cells in lymphoid organs such as bone marrow stromal cells and epithelial cells of the thymus; and in humans it is also expressed in epithelial cells of the skin and intestine^{48,49}. In the thymus, lymphocytes in the earliest stages require IL-7 for survival, proliferation and rearrangement of the TCR (T cell receptor) genes. Mature T cells after leaving the thymus also require IL-7 for survival and homeostatic proliferation. While murine B lymphocytes also require IL-7 for development, human B cells do not⁴⁹.

IL-7 signals lymphocytes by binding to its specific receptor, IL-7R. The receptor is expressed on the membrane and it is composed of a heterodimer of two transmembrane proteins. The α chain is specific (IL-7R α , also known as CD127) and it is dimerized by a common cytokine receptor γ chain (γ_c), which is also shared by the

receptors of IL-2, IL-4, IL-9 and IL-15. Each of these chains is expressed on the cell surface independently of each other, although the two chains could pre-associate. The subunits as monomers or as homodimers are not sufficient to bind IL-7; both of these subunits are essential for efficient signaling⁴⁹.

The human IL-7R α gene is mapped on chromosome 5 with a size of ~20kb while the murine IL-7R α gene is on chromosome 15 with a size of ~22kb. Both human and murine genes contain eight exons and seven introns. IL-7R is composed of 439 amino acids in its mature form with a molecular weight of 49.5 kD. IL-7R α is mainly expressed in cells of the lymphoid lineage, such as T lymphocytes, progenitor B-lymphocytes, NK cells, dendritic precursors, and bone marrow-derived macrophages. Although their role in non-lymphoid cells is not well known, IL-7R α is also expressed in normal human intestinal epithelial cells, endothelial cells, colorectal cancer cells, breast cancer cells and some other malignant cell lines^{49,50}.

1.2.2 IL-7 Receptor Signaling Pathways

The components of the IL-7 receptor, IL-7R α chain and common γ chain dimerize upon IL-7 extracellular cytokine binding. Dimerization activates kinases bound to γc and IL-7R α on the intracellular domains, JAK3 and JAK1. The phosphorylation of IL-7R α intracellular domain by JAK1 triggers recruitment of PI3K and STAT proteins. JAK protein phosphorylates STAT proteins, which results in their dimerization and translocation to cell nucleus in order to bind transcription activation sites of genes such as Bcl-2, SOCS-1, cyclinD1 and c-myc to promote cell differentiation and survival. Also, recruitment of PI3K to the IL-7R α intracellular domain activates this kinase and results in phosphorylation of the Akt, which promotes cell survival by causing the degradation of pro-apoptotic proteins such as Bad and Bax (Figure 1.7)⁵¹.

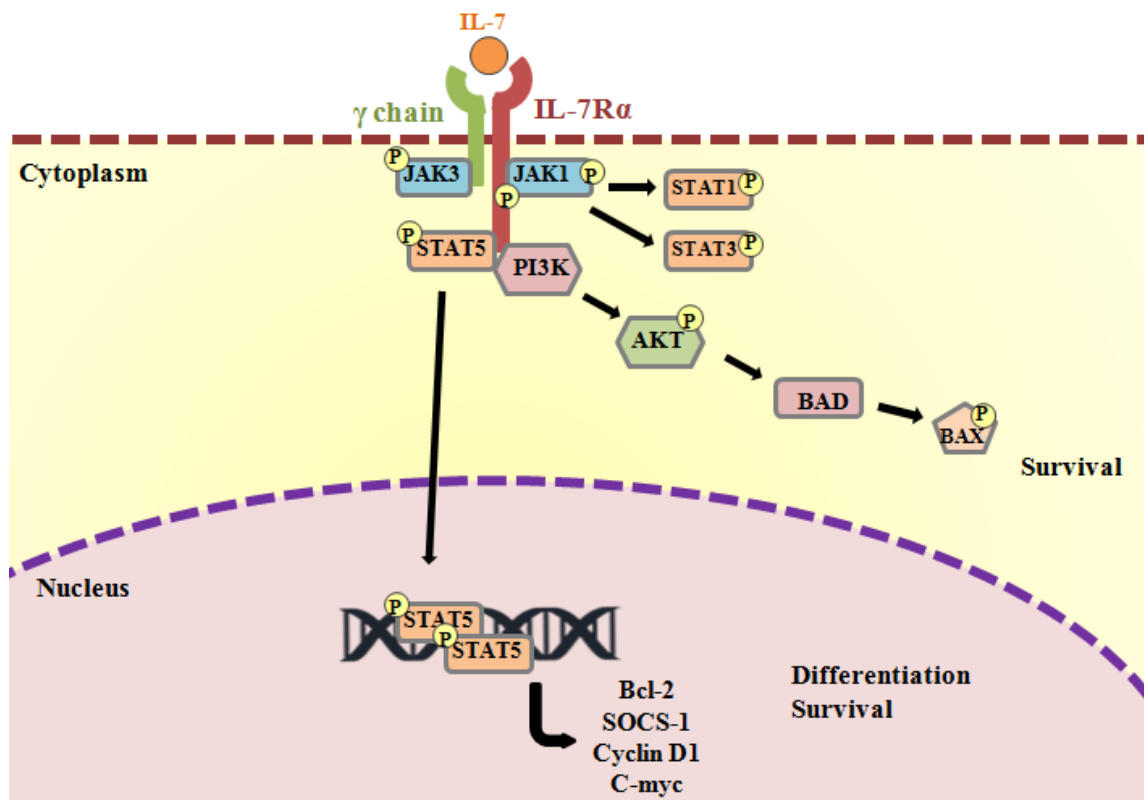


Figure 1.7 The IL-7 receptor signaling pathway ⁵¹.

1.2.3 Importance of the IL-7R Signaling for Lymphopoiesis

B cell development mainly takes place in the bone marrow and it can be subdivided into various stages depending on the expression of different intracellular and cell surface markers, the rearrangement status of the heavy and light immunoglobulin chains, and their cell cycle status ⁵². The most important stages of B cell development in the bone marrow and their IL-7R expression pattern are shown in Figure 1.8.

The block of transition from the pro-B cell to pre-B cell stages in IL-7R deficient mice indicates that IL-7 signaling has an essential role in B lymphocyte development ⁵³. Later on it was reported that IL-7Rα regulates access to immunoglobulin heavy chain coding gene segments during somatic recombination, which is a critical step for the diversity of antibodies ⁵⁴. In the pre-B cell stage, IL-7R signaling is stopped before the rearrangement of the light chain gene locus by the upregulation of the IRF-4 transcription factor ⁵⁵. The transition from the pro-B stages to

later ones is regulated by transcription factors such as EBF, which is upregulated by IL-7R signaling⁵⁶.

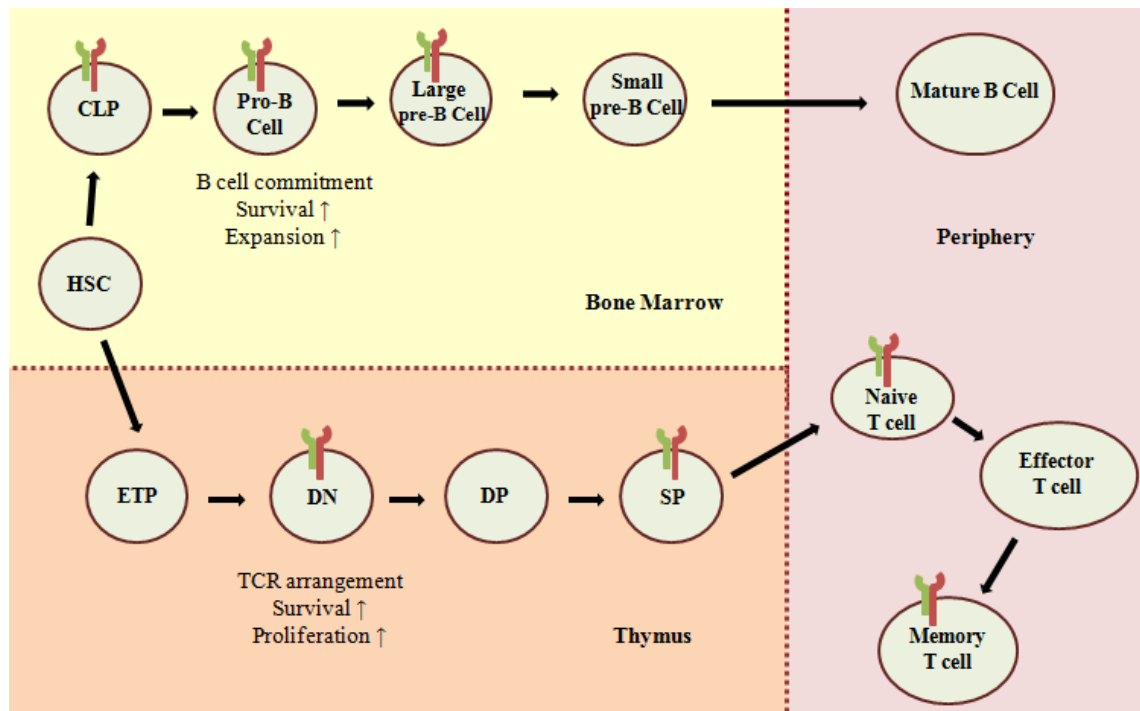


Figure 1.8 IL-7R expression by lymphocytes⁵⁷. B lymphocytes of the bone marrow and T lymphocytes of the thymus express IL-7R on the cell surface at different stages of development. The expression of IL-7R is dynamically regulated during development.

IL-7 signaling is essential for T lymphocytes throughout their life cycle; it has roles in T cell maturation, differentiation and mature T cell survival in peripheral lymphoid tissues. IL-7R expression is strictly regulated over the course of T cell development. It is expressed on double-negative thymocytes, its expression is turned off in the double positive stage and it is re-expressed in single-positive stage (Figure 1.8). IL-7 signaling first occurs at the double negative (DN) stage which can be subdivided into four groups from DN1 to DN4, determined by the surface expression of CD44 and CD25. IL-7R α expression starts at the DN2 stage and the signaling occurs at the DN3 stage, where TCR β selection of the DN thymocytes occurs. IL-7R signaling is critical for the survival and proliferation of these selected cells. IL-7 deficient cells are normally found to be developmentally arrested at the DN3 stage and overexpression of anti-apoptotic molecules such as Bcl-2 or deletion of pro-apoptotic factors such as Bim and Bax can compensate for the lack of IL-7 signaling which shows that in this stage IL-7 is responsible of providing survival signals (reviewed in⁵⁸). IL-7R expression is downregulated in the DN4 stage and terminated in immature CD4+CD8+ double

positive (DP) thymocytes. This termination is also important since IL-7 signaling inhibits expression of the transcription factors TCF-1 and LEF-1, which are essential for DP cell differentiation⁵⁹. DP thymocytes are metabolically inactive and pre-programmed for cell death. As a result of positive selection, post DP intermediate cells ($CD4^+CD8^{low}$) start to express IL-7R again, and CD8 coreceptor transcription is downregulated. Persisting TCR signaling differentiates intermediate cells into CD4 SP cells, whereas intermediate cells that no longer receive TCR signals differentiate into CD8 SP cells due to IL-7 signaling⁶⁰.

IL-7 is also a central regulator of peripheral T-cell homeostasis and survival for both naive and memory CD4 and CD8 T cells. IL-7R is only down-regulated upon T cell activation and memory cell differentiation when other γ chain cytokines such as IL-2 and IL-15 take over survival signaling. After differentiation is complete IL-7R is re-expressed in memory T cells⁶¹.

1.2.4 Regulation of the IL-7R alpha Gene

During the development of both B and T lymphocytes IL-7R α is expressed differentially and its expression is strictly regulated throughout their life cycle. The expression in different stages is controlled by various transcription factors. Bioinformatically identified transcription factor binding sites on the IL-7R α gene locus are shown in Figure 1.9.

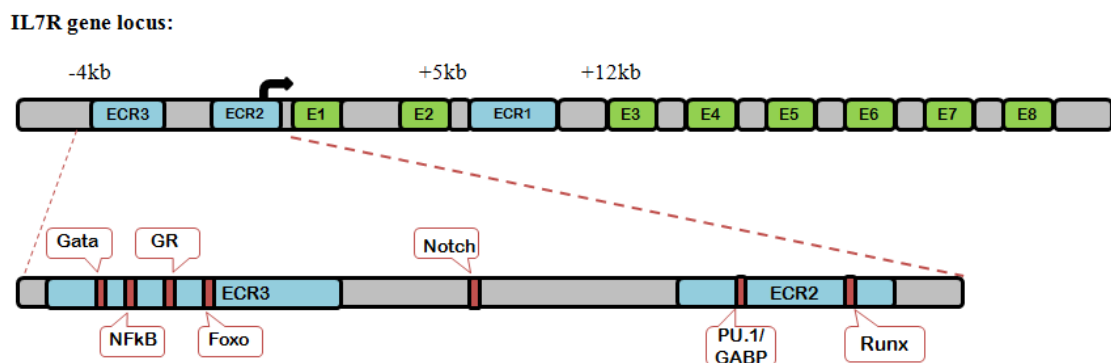


Figure 1.9 IL-7R gene locus with various transcription factor binding sites.

In the gene locus there are three evolutionarily conserved regions (ECR); one in between the second and the third exons, one in the promoter region, and one in the upstream region. In the IL-7R α promoter there is a GGAA motif region which is a binding site for the PU.1 transcription factor. PU.1 is an ETS family transcription factor and was demonstrated to take role in IL-7R expression of developing B cells ⁶². In T cells this site is bound by GGAA binding protein (GABP), which is also an ETS family transcription factor. In the absence of PU.1 GABP can promote IL-7R expression in committed B cells too, but not in early B cell progenitors ⁶³. Another transcription factor that binds the IL-7R promoter region is Runx1. The deficiency of Runx1 in CD4 T cells was shown to result in reduced IL-7R expression levels and a shorter survival period. This indicated that Runx1 transcription factor was necessary for the positive selection and maturation of CD4 SP cells by IL-7 mediated survival signaling ⁶⁴.

About 3 kb upstream of the IL-7R α transcription initiation site there is another evolutionarily conserved region which contains the binding sites for transcription factors Gata, NFkB, glucocorticoid receptor (GR) and Foxo. GATA-3 is a zinc-finger transcription factor that is essential for the generation of the earliest T cell progenitors and it was shown that GATA-3 expression was required for the generation of IL-7R positive thymus derived NK cells while bone marrow derived NK cells could develop in its absence ⁶⁵. Foxo1 is a transcription factor that has various roles in cell regulation. In T cells it was shown that Foxo1 deficiency resulted in severe defect in IL-7R α expression. Other factors that have roles in the regulation of naïve T cell homeostasis and life-span were also shown to be affected by Foxo1 ⁶⁶. NFkB, which is another transcription factor that has a binding site in the IL-7R enhancer region, is an important regulator for the activation, proliferation and survival of thymocytes. It was recently shown that IL-7 stimulation on NFkB deficient T cells did not enhance their viability and those cells appeared to have reduced IL-7R levels both for the protein and mRNA; demonstrating the role of NFkB in the regulation of IL-7R expression ⁶⁷.

1.2.4.1 Notch Transcription Factor

The Notch pathway regulates cell proliferation, cell fate, differentiation, and cell death in all metazoans. Notch is a cell-surface receptor that sends short-range signals by interacting with transmembrane ligands such as Delta (Delta-like in humans) and

Serrate (Jagged in humans) on neighboring cells. Upon ligand binding, the ligand-receptor complex unfolds a juxtamembrane negative control region which allows access of ADAM10 protease to site 2 (S2) to cleave the Notch extracellular domain. Then γ -secretase induces the second cleavage at site 3 (S3) to release the Notch intracellular domain (NICD) from the membrane. NICD then travels to nucleus to bind transcriptional complexes containing a DNA binding protein CBF1/RBPjK/Su(H)/Lag1 (CSL) and its co-activator Mastermind-like (MAML) to regulate target gene expression (Figure 1.10) ⁶⁸.

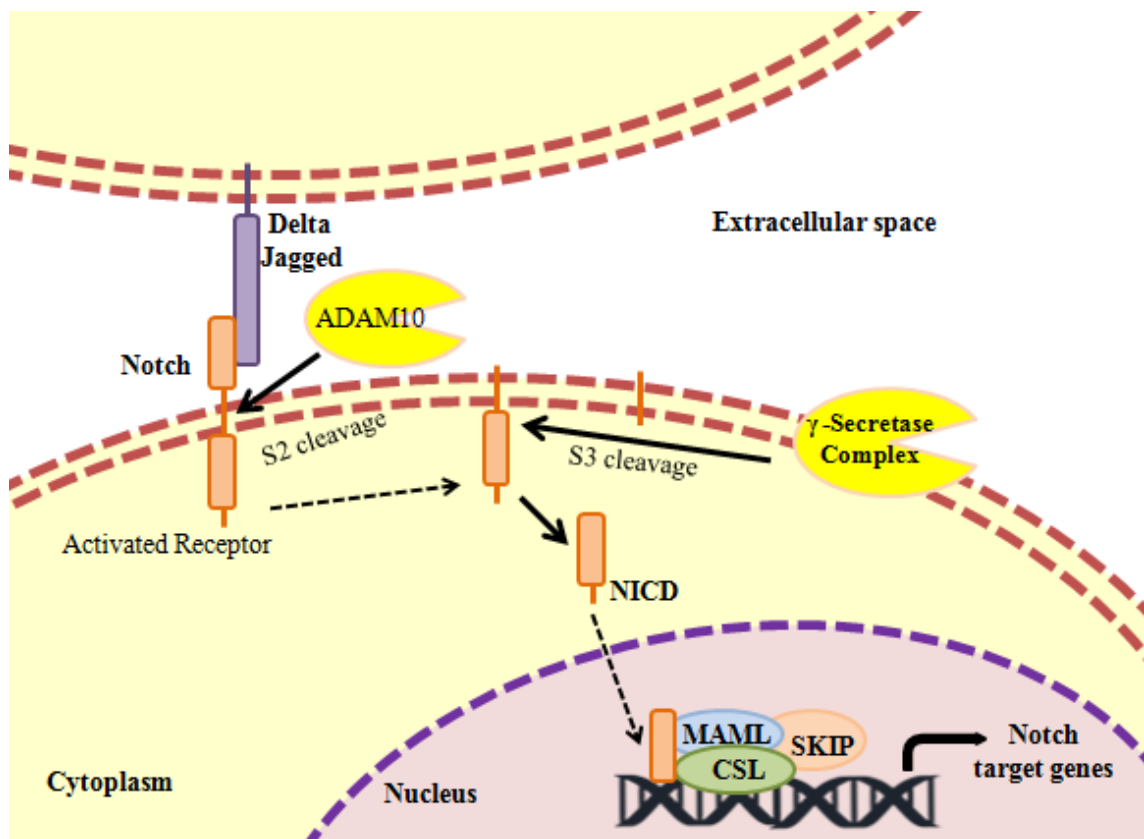


Figure 1.10 Notch signaling ⁶⁸. The Notch intracellular domain (NICD) is released from the membrane upon ligand binding induced cleavage of the Notch receptor on the plasma membrane. Cleaved NICD translocates into the nucleus, binds a preexisting CSL (RBP-Jk) transcription factor complex, helps recruit of the adaptor protein Mastermind-like (MAML) and promotes transcriptional activation.

The studies done with mice have shown that the loss of Notch1 could cause a reduction in thymus size and deficiency in thymocyte development. In addition, while Notch1-deficient bone marrow can contribute to the development of all hematopoietic cells normally, T cell development was blocked at an early stage, before the expression

of T cell lineage markers ⁶⁹. Also in another study Notch1 was found to bind CSL binding motif (RBP-Jk) IL-7R *in vivo* and in early thymic precursors Notch1 ectopic expression consistently resulted in the generation of DP thymocytes with up-regulated IL-7R while defective Notch1 signaling impaired their IL-7R expression, resulting in a developmental arrest that could be rescued by ectopic expression of IL-7R ⁷⁰. These findings suggest that Notch1 has an essential and selective role in T cell maturation.

1.2.4.2 Glucocorticoid Receptor (GR)

Activated lymphocytes or macrophages secrete inflammatory cytokines such as TNF- α and IL-1 β to activate components of the inflammatory system. As a result of an endocrine feedback loop, the release of these cytokines subsequently stimulate the cells of the adrenal cortex and they secrete glucocorticoids (GC) to induce anti-inflammatory effects on immune cells through interruption of proinflammatory cytokine-mediated signaling pathways or by apoptosis ⁷¹. Along with these, GCs are shown to have various effects on the growth, differentiation and function of lymphocytes ⁷².

Due to their lipophilic nature, GCs can readily diffuse through the plasma membrane and bind to the glucocorticoid receptor (GR) in the cytoplasm. GR in its inactive form is in a complex consisting of heat shock proteins (such as hsp90) and a low weight molecular protein p23. GR is activated upon binding its ligand, and it dissociates from its chaperone proteins. Once released, its nuclear localization signals become exposed and it translocates to the nucleus to stimulate or inhibit the activation of its target genes (Figure 1.12). Apart from binding the DNA directly, GR can also induce transcription regulation by binding other factors ⁷¹. GR was shown to bind the IL-7R α upstream enhancer region and activate the IL-7R expression in mouse early stage thymocytes ⁷³. Also, treatment with dexamethasone, which is a synthetic glucocorticoid, was shown to increase IL-7R α expression at both the mRNA and protein levels in mouse and human cells ^{74,75}.

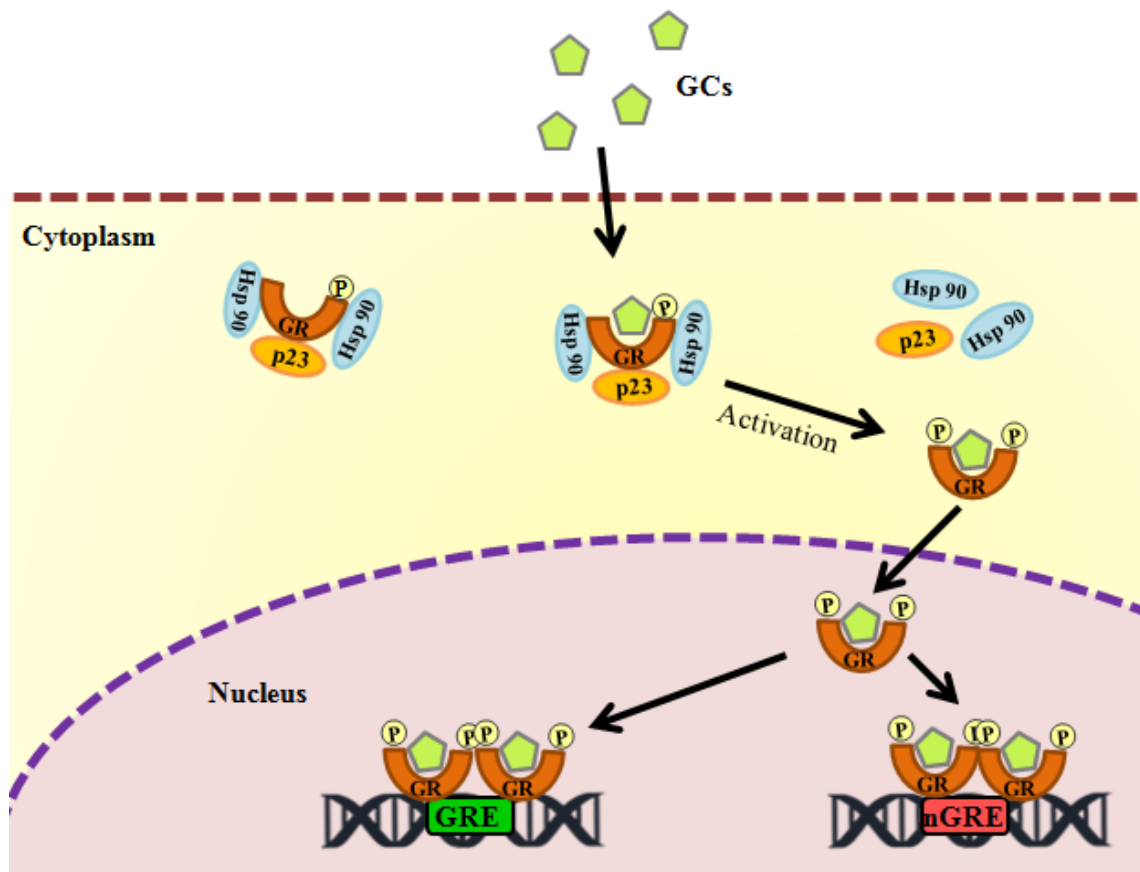


Figure 1.12 Glucocorticoid receptor signaling. Inactive GR is bound by the chaperones Hsp90 and p23 in the cytoplasm until encountering a ligand. Glucocorticoids are hydrophobic and freely diffuse into the cell, and upon binding GR it causes its activation. Active GR dissociates from its chaperones, is phosphorylated and translocates into the nucleus to bind DNA. GR can activate genes that have glucocorticoid response elements (GRE) in their promoter and inhibit genes that have negative GRE (nGRE) ⁷¹.

2. AIM OF THE STUDY

Transcription activator-like effector (TALE) proteins from the plant pathogen *Xanthomonas* consist of highly conserved repeat units in their central DNA binding domain (DBD). Binding specificity of the each repeat is determined by the polymorphic amino acid residues at positions 12 and 13, named as repeat variable di-residues (RVDs). The simplicity of one RVD - one base code and the modular structure of the DBD enable the assembly of proteins that can target any site in the genome with high specificity and various functions. TALE nucleases (TALENs), which are generated by the fusion of the non-sequence specific DNA cleavage domain of FokI to the TALE DNA binding domain, have become an important tool for targeted gene modification. TALENs induce double strand breaks (DSBs) at the targeted site and cellular repair of the disruption occurs either by non-homologous end joining (NHEJ) or by homologous recombination (HR) which results in site directed mutagenesis. We used TALENs to modify the genome of mouse cells.

In the first part of the study, we targeted the IL-7R α gene and aimed to mutate two of the transcription factor binding sites in the upstream enhancer region to observe the effects of these mutations on IL-7R expression. We also targeted the exon 2 and exon 3 of IL-7R to delete a transcriptional control element in intron 2. We designed and constructed TALEN pairs targeting the binding sites of glucocorticoid receptor (GR) and Notch transcription factors, along with the exon targeting ones and expressed them in the murine RLM11 cells. We used the restriction fragment length polymorphism (RFLP) assay and DNA sequencing to detect mutations and we monitored IL-7R expression by flow cytometry. In the second part of the study we targeted the gene encoding the GR transcription factor and designed a TALEN pair against the translation start site to knockout the gene. Also, we designed a donor plasmid homologous to the same site with a Venus-YFP gene insertion to fuse endogenous GR with Venus through homologous recombination. With the knockout experiment, we aimed to observe the effects of GR gene deficiency and with homologous recombination we aimed to generate a model that would enable tracking GR activities within the cell.

3. MATERIALS AND METHODS

3.1 Materials

3.1.1 Chemicals

All the chemicals used in this project are listed in the Appendix A.

3.1.2 Equipment

All the equipment used in this project are listed in the Appendix B.

3.1.3 Buffers and Solutions

Standard buffers and solutions used in this project were prepared according to the protocols in Sambrook et al., 2001.

Calcium Chloride (CaCl₂) solution: 60 mM CaCl₂, 15% glycerol and 10mM PIPES at pH 7.00 were mixed and the solution was filter-sterilized and stored at 4°C for competent cell preparation.

5X Tris-Borate-EDTA (TBE) Buffer: 54 g Tris base, 27.5 g Boric acid and 20mL of 0.5 M EDTA at pH 8.00 were dissolved in 1L of dH₂O and stored at RT.

1% (w/v) Agarose gel: 1 g of agarose was dissolved in 100 mL of 0.5X TBE buffer by heating in a microwave oven. 0.001 % (v/v) of ethidium bromide was added to the solution for visualization of nucleic acids.

Phosphate-buffered saline (PBS): 1 tablet of PBS was dissolved in 200 mL of dH₂O. The solution was filter-sterilized for use in mammalian cell culture and stored at 4°C.

Polyethylenimine (PEI) (1µg/µL): 50 mg PEI was dissolved in 50 mL dH₂O that has been heated to ~80°C and cooled to room temperature. After neutralizing to pH 7.00, the solution was filter-sterilized, aliquoted and stored at -20°C.

FACS buffer: 0.5 g Bovine serum albumin (BSA) and 0.5 g sodium azide were dissolved in 500 mL 1X HBSS and stored at 4°C.

3.1.4 Growth Media

3.1.4.1 Bacterial growth media

Liquid media: 20 g Luria-Broth (LB) was dissolved in 1 L of dH₂O and autoclaved at 121°C for 15 min. For selection, ampicillin with a final concentration of 100 µg/mL and spectinomycin with a final concentration of 50µg/mL were added to liquid medium after autoclave.

Solid media: 35 g LB agar was dissolved in 1 L of dH₂O and autoclaved at 121°C for 15 min. For selection, antibiotics with previously indicated concentrations were added to autoclaved medium after cooling down to 50°C. Autoclaved and antibiotic added medium was poured onto sterile Petri dishes. Solid agar plates were stored at 4°C.

3.1.4.2 Mammalian cell culture growth media

The adherent cell line HCT116 was grown in DMEM cell culture medium that was supplemented with 10% heat inactivated fetal bovine serum (FBS), 2 mM L-glutamine, 100 unit/mL penicillin and 100 unit/mL streptomycin.

The suspension cell line RLM11 was grown in RPMI 1640 cell culture medium that was supplemented with 10% heat inactivated fetal bovine serum (FBS), 2 mM L-

Glutamine, 100 unit/mL, 100 unit/mL streptomycin, non-essential amino acids, vitamin and 50 μ M 2-mercaptoethanol.

Both adherent and suspension cell lines were frozen in fetal bovine serum (FBS) containing DMSO at a final concentration of 10% (v/v). Freezing medium was stored at 4°C.

3.1.5 Cell Types

E. coli DH-5 α competent cells were used for bacterial transformation of plasmids. RLM11, a radiation-induced BALB/c murine CD4 single positive thymoma T cell line, was used for transfection and analysis of IL7R expression level with FACS ⁷⁶. HCT116, a human colonic carcinoma cell line that is generally used for tumorigenicity studies, was the only adherent cell line in this study and was used for GR TALEN project.

3.1.6 Commercial Molecular Biology Kits

- QIAGEN Plasmid Midi Kit, 12145, QIAGEN, Germany
- QIAquick Gel Extraction Kit, 28704, QIAGEN, Germany
- GenElute Mammalian Genomic DNA Miniprep Kit, G1N350, SIGMA, Germany
- GenElute PCR Clean-Up Kit, NA1020, SIGMA, Germany
- CloneJET™ PCR Cloning Kit, K1232, Thermo Fisher Scientific.
- InsTAclone PCR Cloning Kit, K1214, Thermo Fisher Scientific
- Gibson Assembly® Master Mix, New England BioLabs, UK

3.1.7 Enzymes

All enzymes and their corresponding buffers used in this project are from NEB and Fermentas.

3.1.8 Vectors and Primers

Vectors and primers used in this project are listed in Table 3.1 and Table 3.2.

Vector Name	Purpose	Bacterial Resistance
pcDNA-GFP	Transfection efficiency control	Ampicillin
pUC19	Transformation efficiency control	Ampicillin
pHD1-pHD10 pNG1-pNG10 pNH1-pNH10 pNN1-pNN10 pNI1-pNI10	Module plasmids for TALE / TALEN construction	Tetracycline
pLR-HD pLR-NG pLR-NH pLR-NN pLR-NI	Last repeat plasmids for TALE / TALEN construction	Tetracycline
pFUS_A pFUS_B1- pFUSB10	Array plasmids for TALE / TALEN construction	Spectinomycin
pC-Goldy TALEN	Backbone plasmid for TALEN construction	Ampicillin
pJET1.2/blunt	Cloning of PCR products	Ampicillin
pTZ57R/T	Cloning of PCR products	Ampicillin
hAAVS-SA2A-1	Donor plasmid for Puromycin resistance gene	Ampicillin
mVenus-C1	Donor plasmid for Venus-YFP gene	Kanamycin

Table 3.1 List of vectors used in this project

Primer Name	Sequence	Purpose
Notch for BamHI	ATAGGATCCATTGAAACCATAACCACCCTC	Notch TALEN Target site amplification
Notch rev Bgl2	GCGAGATCTCCCTTCTCTCTAATTCTGTT	Notch TALEN target site amplification
Kpl11 For	CCAAGGAATAAACCCAAGGA	IL7R upstream region amplification
Kpl12 Rev	AGAAGCACGCTTGTATGTGC	IL7R upstream region amplification
hGRTalenFwd	AGCTTATGATGTTTTCCCCCGTTTTTG	hGR TALEN target site amplification
hGRTalenRev	AGTCCATCACATCTCCCCTCTCCT	hGR TALEN target site amplification
nfkb TALEN for	CTTCCCGCACTCTATTAGAT	IL7R-GR TALEN target site amplification
nfkb TALEN rev	CTTTCATGGGCTATCACTCC	IL7R-GR TALEN target site amplification
Int2R1 Fwd	CCTTCATGTCTGCCACTCAA	IL7R Exon 2 TALEN target site amplification
Int2R1 Rev	CATATTTGAAATTCCAGATTAGCTGT	IL7R Exon 2 TALEN target site amplification
Int2R2 Fwd	TGGGGCTCTTTTACGAGTG	IL7R Exon 3 TALEN target site amplification

Int2R2 Rev	GCAAAAATAGTTGCTCATGTTTATT	IL7R Exon 3 TALEN target site amplification
pCR8_F1	TTGATGCCTGGCAGTTCCCT	Colony PCR of Golden GATE reaction #1
pCR8_R1	CGAACCGAACAGGCTTATGT	Colony PCR of Golden GATE reaction #1
TAL_F1	TTGGCGTCGGCAAACAGTGG	Colony PCR of Golden GATE reaction #2
TAL_R2	GGCGACGAGGTGGTCGTTGG	Colony PCR of Golden GATE reaction #2
SeqTALEN_5-1	CATCGCGCAATGCACTGAC	Sequencing of final TALEN construct
pJET1.2 forward sequencing primer	CGACTCACTATAGGGAGAGCGCC	Colony PCR and sequencing of cloned PCR products
pJET1.2 reverse sequencing primer	TTCTTGTAGCTAAAAGGTACCGTC	Colony PCR and sequencing of cloned PCR products
hGRLeftArm Fwd	TGGCTAGCGTCTGTCTCGGAAGATAAGCAGA TCAGCATTGTTTA	hGR homologous recombination donor construct
hGRLeftArm Rev	CAGTGAATATCAACTACAAAACAAAAAAC AAAAACGGG	hGR homologous recombination donor construct
hgrPuro Fwd	TTGTAGTTGATATTCAGTATGACCGAGTA CAAGCCCACGGTGC	hGR homologous recombination donor construct

hgrPuro_P2A Rev	ACGTCTCCTGCTTGCTTTAACAGAGAGAAG TTCGTGGCGGCACCGGGCTTGCGGGTC	hGR homologous recombination donor construct
hgrP2A_Venus fwd	TAAAGCAAGCAGGAGACGTGGAAGAAAAC CCCGGTCCCATGGTGAGCAAGGGCGAGGA GCT	hGR homologous recombination donor construct
hgrVenus Rev	AGCTCGAGATCTGAGTCCGGACTTGTACAG	hGR homologous recombination donor construct
hGRRightArm Fwd	CTCAGATCTCGAGCTATGGACTCCAAAGAA TCATTAACCTCGGTAGAG	hGR homologous recombination donor construct
hGRRightArm Rev	GTGGATCCGACTCCAAATCCTGCAAAATGT CAAAGGTGC	hGR homologous recombination donor construct

Table 3.2 List of primers used in this project

3.1.9 DNA Molecular Weight Marker

DNA molecular weight marker used in this project is given in Appendix C.

3.1.10 DNA sequencing

DNA sequencing was commercially performed by McLab, CA, USA.

(<http://www.mclab.com/home.php>)

3.1.11 Software and Computer Based Programs

The software and computer based programs used in this project:

Program Name	Website/ Company	Purpose
CLC Main Workbench 6.1.1	http://www.clcbio.com/	Primer design, molecular cloning, sequence data management
FlowJo 7.6.5	http://www.flowjo.com/	FACS data analysis

TAL Effector Nucleotide Targeter 2.0	https://tale-nt.cac.cornell.edu/	TALE / TALEN design tool
Quantity One	Bio – Rad	Gel image analysis
Visual Molecular Dynamics (VMD)	http://www.ks.uiuc.edu/Research/vmd/	Crystal structure display and analysis

Table 3.3 List of software and computer based programs used in this study

3.2 Methods

3.2.1 Bacterial Cell Culture

3.2.1.1 Bacterial culture growth

E. coli DH5 α bacterial cells were grown overnight (~16 h) at 37°C shaking at 250 rpm in Luria Broth (LB). Bacterial cells were either spread or streaked on LB Agar plates to obtain single colonies and grown overnight (~16 h) at 37°C. Antibiotics were added to growth media depending on the application. For long-term storage of bacterial cells, glycerol was added to the overnight grown culture to a final concentration of 15% in 1 mL. Bacterial glycerol stocks were stored at -80°C.

3.2.1.2 Competent cell preparation and transformation

E. coli DH5 α competent cells were prepared using stock of previously prepared competent cells. 50 μ L from previously prepared competent cells were grown in 50 mL LB without selective antibiotic overnight at 37°C shaking at 250 rpm. Next day, 4 mL from the overnight culture was diluted within 400 mL LB and incubated under same growth conditions until the OD₅₉₀ reaches to 0.375. Then, previously prepared ice-cold CaCl₂ solution was used for resuspension of bacterial cell pellet after successive centrifugation steps and for final preparation. 200 μ L aliquots of competent cells

prepared were frozen immediately in liquid nitrogen and then stored at -80°C. Competency of prepared cells was tested by transforming varying concentrations of pUC19 plasmid.

For transformation of competent cells, CaCl₂ treated chemically competent bacterial cells were taken from -80°C and ~100 pg of plasmid DNA was added before cells were completely thawed. After 30 min of incubation on ice, the cells were heatshocked at 42°C for 90 seconds and transferred back to ice rapidly to chill for 60 seconds. 800µL of sterile LB without antibiotics added and cultures were incubated for 45 minutes at 37°C for recovery of cells and expression of antibiotic resistance gene encoded by the plasmid. Transformed cells were spread onto LB agar plates containing appropriate antibiotic for selection using sterile glass beads. Then, the plates were incubated overnight at 37°C.

3.2.1.3 Plasmid DNA isolation

Plasmid DNA isolation was performed either by the alkaline lysis protocol or by QIAGEN Plasmid Midi Kits. For plasmid isolation, either a single colony of *E.coli* from LB agar plates or a stab of the glycerol stock was grown overnight at 37°C shaking at 250 rpm in liquid medium containing selective antibiotics with appropriate concentrations. The concentration and purity of kit-isolated plasmid DNA were determined using Nanodrop.

3.2.2 Vector Construction

3.2.2.1 The General Methods Used in Vector Construction

Restriction Enzyme Digestion: Digestion reactions containing template DNA, enzyme and its compatible buffer were incubated at the optimum temperature of the enzyme used for 2 hours. ~ 300 ng of template DNA was used for diagnostic digestions whereas the amount of template DNA used for gel extraction and cloning purposes was at least 1 µg.

Agarose Gel Electrophoresis and Gel Extraction: Agarose gels to observe DNA samples and digestion products were prepared in varying concentrations from 1% to 2% depending on the size of DNA fragments to be separated. Agarose gel was prepared by dissolving an appropriate amount of agarose powder in 0.5X TBE, heating for 3-5 min in a microwave. After cooling-down of the solution to room temperature, ethidium bromide was added at a final concentration of 0.001% (v/v) and the gel was poured onto the gel apparatus for solidification. 0.5X TBE was also used as running buffer. DNA samples were mixed with 6X DNA loading dye before loading to the gel. Electrophoresis was performed at 100-135 V for 45-75 minutes and the bands were observed under UV light. Gel extraction of DNA samples was performed using a QIAGEN Gel Extraction Kit.

Dephosphorylation of Vector Ends: 5' phosphate groups of linearized vector DNA were dephosphorylated using Calf Intestinal Alkaline Phosphatase (CIAP) prior to insert ligation, to prevent vector re-ligation.

Ligation: Ligation was performed using T4 DNA Ligase (Fermentas), in 1:3 and 1:6 molar vector to insert ratio using 100ng vector. In addition, ligation reaction mixture without insert was prepared as negative control for each ligation. The ligation reaction was incubated at 16°C for 16 hours in a final volume of 20µL. Then, half of the ligation mixture was transformed into chemically competent bacteria.

3.2.2.2 Vector Construction for Homologous Recombination

Homologous sites to the targeted region as well as insertion sites that were obtained from other plasmids were separately PCR amplified before being fused with other PCR products and cloned into their final vector. In order to fuse PCR products with each other in the desired order, primers were designed to have homologous overhangs at their 5' ends. After PCR, the correct bands were gel extracted and purified from their respective primers; and two or more of these PCR amplicons were used as a template in another polymerase chain reaction. In the second PCR amplification with multiple templates only the primers in the 5' and 3' end of the first two PCR amplicons

were used. Because such reactions often generate more than one band, the correct band was gel extracted again before further use. The primers at the either end had restriction enzyme cut sites instead of a homologous overhang to enable restriction digestion and ligation of the final product into a destination vector.

This method was used to fuse exon 2 and exon 3 of IL7R gene to use it as a donor⁷⁷ and also to construct hGR gene homologous plasmid with Venus-YFP and puromycin resistance insertions. The strategies used for these are summarized in Figure 3.1 and Figure 3.2 in the respective order. The PCR conditions for the hGR donor plasmid construction are given in Table 3.4, Table 3.5 and Table 3.6. After obtaining the desired construct the gel extracted DNA was either digested on both ends and cloned into a plasmid that had the same restriction enzyme cut sites or it was cloned into the pJET1.2/blunt plasmid without any digestion using the CloneJETTM PCR Cloning Kit (Thermo Scientific).

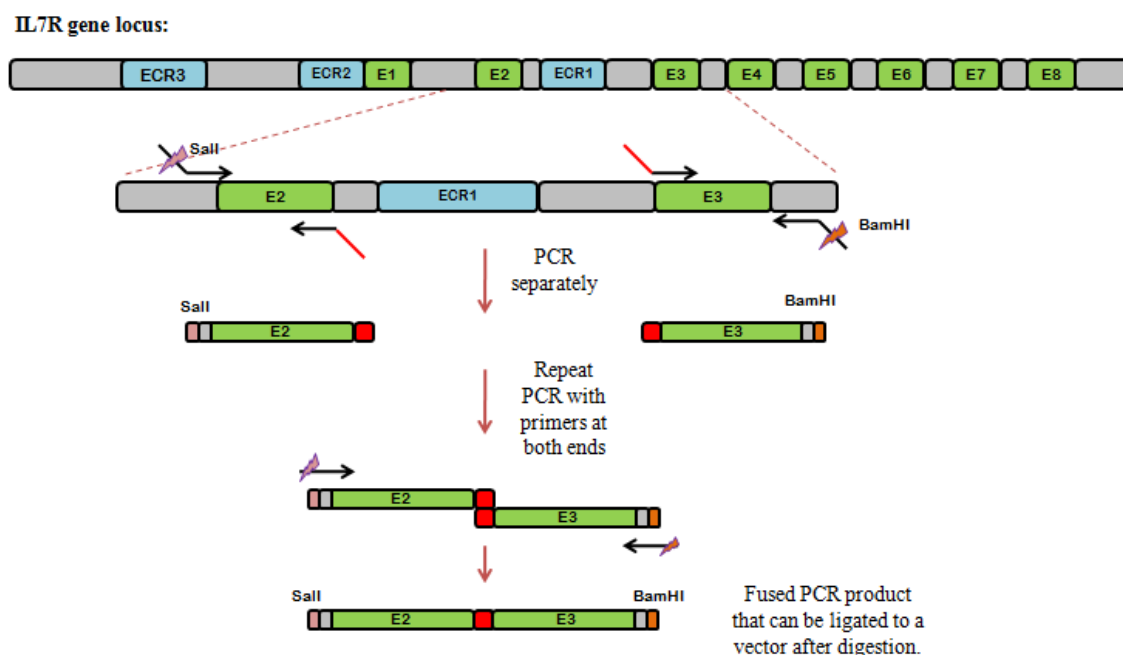


Figure 3.1 The strategy for fusion of exon 2 and exon 3 of IL7R gene.

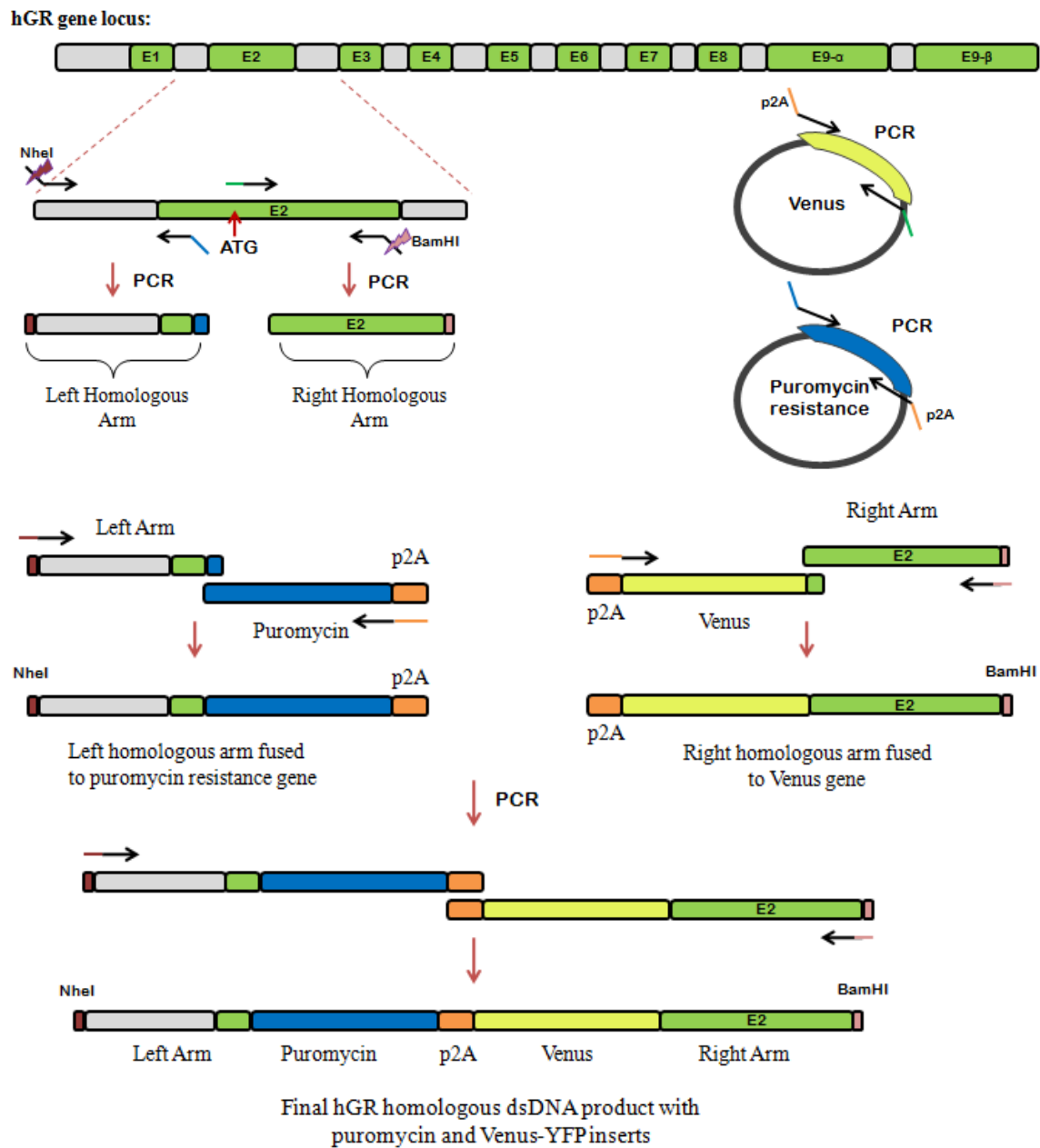


Figure 3.2 The strategy for construction of puromycin resistance and Venus-YFP inserted hGR gene homologous plasmid.

Component	Amount
Template genomic DNA or template Plasmid/Vector (10ng/μl)	1.0μL
5X Phusion high fidelity buffer	5.0μL
10mM dNTP each	0.4μL
Forward primer (10mM)	1.0μL
Reverse primer (10mM)	1.0μL
Phusion Hot Start II DNA polymerase (2U/μl)	0.2μL
dH ₂ O	16.4μL
Total	25μL

Table 3.4 The components and the amounts they were used for separate PCR reactions that were done in hGR donor plasmid construction.

PCR was performed according to the following cycle:

For Left Arm and Right Arm:

98°C/4 min + 30X (98°C/30 s + 62°C/30 s + 72°C/60 s) + 72°C/10 min

For Venus:

98°C/4 min + 30X (98°C/30 s + 63°C/30 s + 72°C/60 s) + 72°C/10 min

For Puromycin resistance:

98°C/4 min + 30X (98°C/30 s + 72°C/60 s) + 72°C/10 min (2-step)

Component	Amount
Gel extracted PCR product (10ng/μl)	1.0μL each
5X Phusion high fidelity buffer	5.0μL
10mM dNTP each	0.5μL
Forward primer (10mM)	1.0μL
Reverse primer (10mM)	1.0μL
Phusion Hot Start II DNA polymerase (2U/μl)	0.2μL
dH ₂ O	15.3μL
Total	25μL

Table 3.5 Optimized PCR conditions for Venus-Right Arm fusion.

The PCR was performed according to following cycle:

98°C/4 min + 30X (98°C/30 s + 55°C/30 s + 72°C/2min) + 72°C/10 min

Component	Amount
Gel extracted PCR product (10ng/μl)	1.0μL each
5X Phusion high fidelity buffer	5.0μL
10mM dNTP each	0.5μL
Forward primer (10mM)	2.0μL
Reverse primer (10mM)	2.0μL
Phusion Hot Start II DNA polymerase (2U/μl)	0.2μL
dH ₂ O	13.3μL
Total	25μL

Table 3.6 Optimized PCR conditions for Left Arm - Puromycin resistance gene fusion.

PCR was performed according to following cycle:

98°C/4 min + 31X (98°C/30 s + 50°C/30 s + 72°C/2min) + 72°C/10 min

Gibson Assembly: Another approach for fusing PCR products that have homologous ends is Gibson Assembly (NEB). In the Gibson Assembly Master Mix there are 5' exonuclease, DNA polymerase and DNA ligase that can function at the same temperature. After 5' exonuclease cleaves a portion of the DNA single stranded flanks can anneal with their complementary sequence in another PCR product; DNA polymerase fills in the remaining gaps while DNA ligase seals the nicks. In this project after left homologous arm of hGR gene got fused to puromycin resistance gene and the right homologous arm got fused to Venus-YFP via conventional PCR methods, the final product that combines all four of them was done using Gibson Assembly Master Mix; following the instructions in the kit. Figure 3.3 explains Gibson Assembly's working principle.

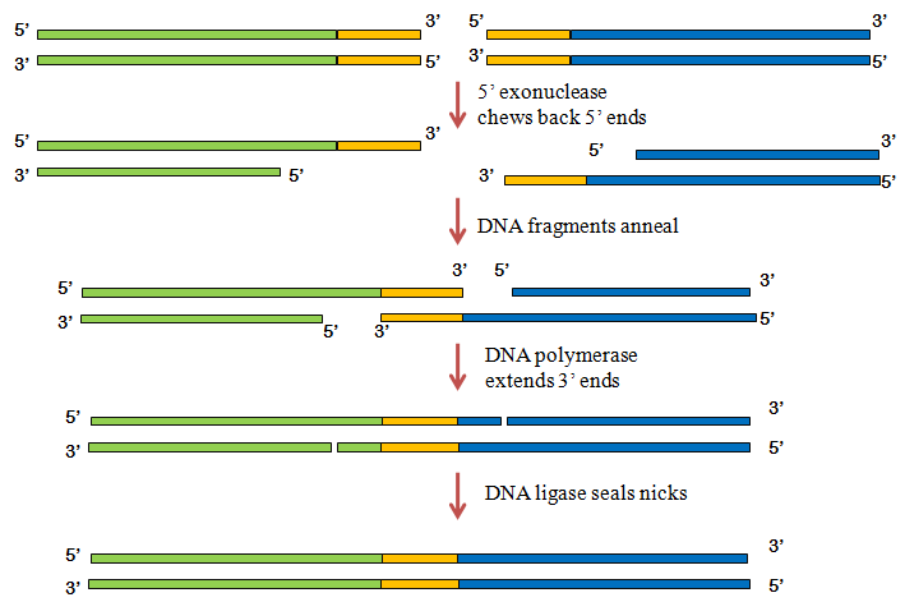


Figure 3.3 Gibson Assembly working principle for PCR products with homologous ends.

3.2.3 Construction of TALEN Expression Vectors

3.2.3.1 Identification TALEN target sites

The software used for design of transcription activator like effectors (TALE) and transcription activator like effector nucleases (TALEN) is available for use as an online tool (TAL Effector Targeter and TALEN Targeter (old version with design guidelines), TALE-NT; <https://tale-nt.cac.cornell.edu/>). DNA sequence entered is scanned for potential TALEN recognition sites based on either preset design guidelines defined by four different articles or user-provided spacer and RVD lengths. The software gives coordinates and sequences of recognition sites for right and left TALEN monomers and the spacer sequence. In addition, RVD sequences necessary for construction of custom TALENs were also provided as software output. Binding sites of TALEN pair and spacer sequences are given in Table 3.7.

	Left TALEN binding sequence	Spacer	Right TALEN binding sequence
hGR TALEN2	CATTAACCTCCTGGTA	gagaagaaaacccca	GCAGTGTGCTTGCTC
Notch TALEN	AGGGTCACCCTCATA	gactcctgggagtttct	ATTGCCCTTGTTTCT
IL7R GR TALEN2	ATTATGTCTTAACTT	tggtcttttacatct	TCACAACTAAAGGAA
IL7R GR TALEN3	ATGTCTTAACTTTGTT	cttttacatcttcaca	ACTAAAGGAAAGAGAT
IL7R Exon2	CACTCCTTCTGGTGCC	acagccagttggaag	TGGATGGAAGTCAACATT
IL7R Exon3	ATATATTTTATAAAGAC	atcagaattcttact	GATTGGTAGCAGCAATAT

Table 3.7 Binding sites of TALEN pair and spacer sequences on the 5'-3' coding strand.

3.2.3.2 Assembly of custom TALEN constructs using Golden Gate TALEN kit

TAL effector DNA binding domain is composed of tandem repeat modules. 12th and 13th amino acids within each repeat module, called repeat-variable di-residues (RVDs), are responsible for nucleotide recognition. NI, NN, NG and HD are the four most common RVDs, each preferentially bind to nucleotides A, G, T, and C, respectively. For some of the TALEN constructs in these experiments modules with NH RVD was used instead of NN to bind nucleotide G. Design of custom TALE and TALENs were performed using TALEN Golden Gate Kit, which was obtained from Addgene. The Golden Gate TALEN kit was reported by Cermak et al (2011) and contains a set of module plasmids with each individual RVDs, array plasmids for intermediate cloning and backbone expression plasmids to make final TALEN expression constructs ¹¹.

The custom TALEN or TAL effector construct is assembled by using successive rounds of Golden Gate cloning, in which digestion by Type IIS restriction endonucleases such as BsaI and Esp3I is performed to create unique 4 bp overhangs on DNA fragments. These unique overhangs flanking each RVD were designed such that up to 10 RVD-encoding repeat module plasmids can be ligated in a single reaction.

Assembly of repeat modules into array plasmids is followed by assembly of array plasmids into final expression vectors (Figure 3.4). Construction of TAL effector or TALEN construct was achieved in 5 days (Figure 3.5).

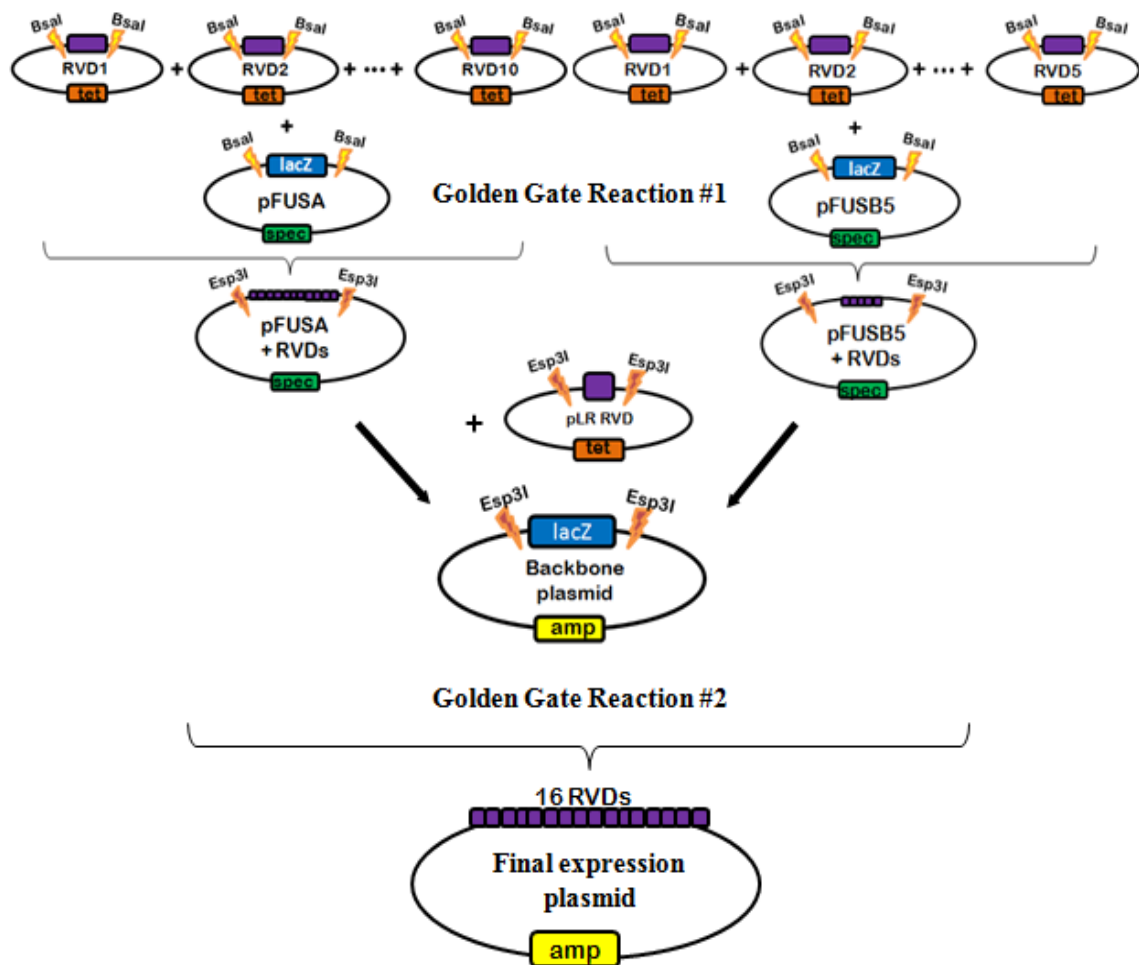


Figure 3.4 Golden Gate assembly of custom TALE and TALEN constructs.

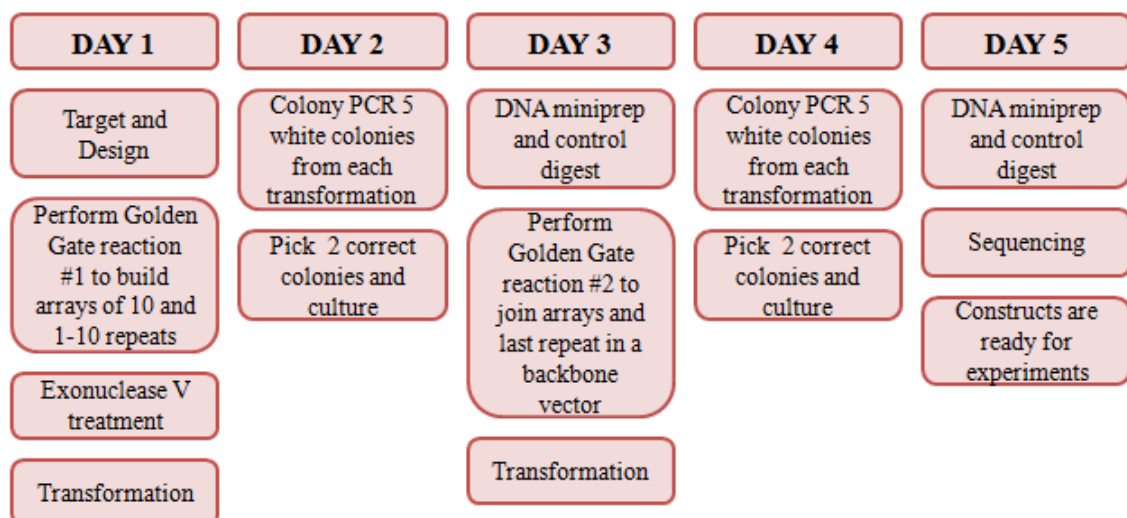


Figure 3.5 Timeline for TALEN construction using TALEN Golden Gate kit ¹¹.

Day 1: After identification of possible TALEN target sites and determination of the RVD sequence, two separate array plasmids were assembled for “n” RVD repeat module containing TALEN expression plasmids. First 10 module plasmids were selected according to the order of RVD sequence, and were cloned into array plasmid pFUS_A. Then, modules selected for remaining RVDs, the ones from 11 to (n-1) were cloned into the array plasmid pFUS_B#n-11. RVD encoding modules for second array were selected starting with plasmid #1. Last RVD (#n) was not included in this reaction as it was provided by a different, “last repeat” plasmid and included in the second step of Golden Gate cloning.

Golden gate reaction #1 was set according to Table 3.8 for each intermediary array plasmid, called as reaction A for first array plasmid and reaction B for the second one. For example a TALEN encoding plasmid with 16 repeats was generated by cloning 10 repeats into the pFUS-A plasmid and 5 repeats into the pFUS-B5 plasmid. The contents of these plasmids were transferred in later days of the procedure into the pCGoldy-TALEN destination expression plasmid along with the contents of last repeat plasmids, in a four plasmid reaction.

Components	Used amount
Each of module vectors	150 ng
pFUS vector	150 ng
BsaI (NEB)	1 μ L
BSA (2 mg/ ml)	1 μ L
T4 DNA ligase (NEB)	1 μ L
10X DNA ligase buffer	2 μ L
dH ₂ O	Up to 20 μ L
Total	20 μ L

Table 3.8 Components and amounts for Golden Gate reaction #1.

Reactions were incubated in a thermo cycler for following cycle:

10 X (37°C/5 min + 16°C/10 min) + 50°C/5 min + 80°C/5min

In order to degrade unligated linear dsDNA fragments of incomplete ligation products, and linearized vectors, 1µL of exonuclease V (RecBCD) (NEB) and 1µL of 10mM ATP were added to reaction and incubated at 37°C for 30min. After completing this procedure, 3.85 µL 100mM EDTA solution was added and the reaction volume was completed to 30 µL with water before performing 70°C incubation for 30min to inactivate exonuclease V and any remaining active enzymes.

Chemically competent DH5α *E.coli* cells were transformed with 2 µL of the reaction and plated on LB agar containing 50µg/mL spectinomycin, with X-gal and IPTG for blue/white screening of colonies.

Day 2: Correct assembly of TALEN RVD repeat modules into intermediary arrays was controlled first by performing colony PCR with 5 white colonies picked from each plate. A PCR master mix was prepared according to colony PCR conditions shown in Table 3.9 using pCR8_F1 and pCR8_R1 as forward and reverse primers, respectively, individual colonies were resuspended in this solution.

Component	Volume
10X standard Taq buffer(Mg free)	2.5µL
25 mM MgCl ₂	2µL
10mM dNTP each	0.5µL
Forward primer	0.2µL
Reverse primer	0.2µL
Taq polymerase (5U/µl)	0.125µL
dH ₂ O	19.475µL
Total	25µL

Table 3.9 Optimized colony PCR conditions.

PCR was performed according to following cycle;

95°C/4 min + 30X (95°C/30 s + 55°C/30 s + 72°C/135 s) + 72°C/10 min

Depending on the colony PCR results, two correct clones were inoculated into 3mL LB containing 50µg/mL spectinomycin and incubated overnight at 37°C shaking at 200 rpm.

Day 3: Plasmid DNA was isolated from overnight cultures of pFUS_A and pFUS_B plasmids containing repeats. Correct assembly of array was controlled by restriction enzyme digestion with AflII and XbaI and agarose gel electrophoresis.

Double digestion with these enzymes releases the repeat arrays and size of fragments produced was 1048 bp for pFUS_A containing 10 RVDs whereas size of fragments varied for pFUS_B plasmids. Correctly assembled intermediary arrays and sequence encoding the n^{th} repeat were assembled into the final expression backbone vector. In these experiments mammalian expression vector pC-Goldy TALEN was used. For Golden Gate reaction# 2, digestion and ligation were performed in 2 steps due to BsmBI restriction enzyme working at 55°C which inhibited the activity of T4 DNA ligase. The first part of the reaction was set according to Table 3.10.

Components	Amount
Reaction A	150 ng
Reaction B	150 ng
pLR vector	150ng
Expression backbone vector	75ng
NEB Buffer 4 (10X)	1.5 μL
BsmBI (NEB)	0.5 μL
dH ₂ O	Up to 15 μL
Total	15 μL

Table 3.10 Components for the first part of Golden Gate reaction #2.

After incubation of the first part of the reaction at 55°C for 10 minutes, the second part of the reaction was set according to Table 3.11.

Components	Used amount
ATP (10 mM)	2 μL
NEB Buffer 4 (10X)	0.5 μL
T4 DNA ligase	1 μL
DTT (0.2M)	1 μL
Water	0.5 μL
Total	20 μL

Table 3.11 Components for second part of Golden Gate reaction #2.

Reactions were incubated in a thermo-cycler using the following cycle:

16°C /15min + 55°C / 15 min + 80°C / 5 min

Chemically competent DH5α *E.coli* cells were transformed with 2 µL of the reaction and plated on LB agar containing 100 µg/mL ampicillin, with X-gal and IPTG for blue/white screening of colonies.

Day 4: Colony PCR was performed to check the assembly of intermediary arrays into the final expression plasmid and 5 white colonies were picked from the plate. The colony PCR mix was prepared according to Table 3, using TAL_F1 and TAL_R2 as forward and reverse primers. After resuspending individual colonies in a reaction mixture, colony PCR was performed according to the following cycle;

95°C/4 min + 30X (95°C/30 s + 55°C/30 s + 72°C/3 min) + 72°C/10 min

Depending on colony PCR results, two correct clones were inoculated into 3mL LB containing 100µg/mL ampicillin and incubated overnight at 37°C shaking at 200 rpm.

Day 5: Plasmid DNA was isolated from overnight cultures and correct assembly of the final full-length repeat array was verified by restriction enzyme digestion with AatII and StuI and agarose gel electrophoresis. In addition, BspEI control digest, which cut only in HD modules of 2-10, was performed to determine final array integrity. DNA midipreps were prepared from correctly assembled TALEN plasmids, and sequenced using the SeqTALEN 5-1 and TAL_R2 primers.

To express TALEN pairs in mammalian cells TALEN constructs should be cloned into a destination plasmid that contains promoter for mammalian expression. In these experiments six different TALEN pairs were designed using Golden Gate TALEN kit; TALEN pairs targeting Notch binding site⁷⁸ and GR binding site of IL7R gene promoter region, TALEN pairs targeting the 2nd and 3rd exons of IL7R gene and finally a TALEN pair targeting the start region of human GR gene. After assembling repeat monomers in array plasmids of pFUS_A and pFUS_B, as destination vector pC-Goldy TALEN backbone was used.

3.2.4 Mammalian Cell Culture

3.2.4.1 Maintenance of mammalian cell lines

In this project a suspension cell line RLM11 was mainly used and it was grown in RPMI 1640 cell culture medium that is supplemented with 10% heat inactivated fetal bovine serum (FBS), 2 mM L-Glutamine, 100 unit/mL, 100 unit/mL streptomycin, non-essential amino acids, vitamin and 50 μ M 2-mercaptoethanol in tissue culture flasks. Additionally adherent cell line HCT116 was used, and it was grown in DMEM cell culture medium that is supplemented with 10% heat inactivated fetal bovine serum (FBS), 2 mM L-glutamine, 100 unit/mL penicillin and 100 unit/mL streptomycin in 10 cm tissue culture dishes. All cultures were maintained in a humidified incubator supplied with 5% CO₂ at 37°C and split into fresh medium when they reach to ~80% confluency. TALEN transfected cells were incubated in a humidified incubator supplied with 5% CO₂ at 32°C after transfection for 72 hours.

For preparation of frozen stocks of both adherent and suspension cell lines, cells at exponential growth phase were resuspended in ice-cold freezing medium. They were stored at -80°C for 24-48 hours and then transferred to liquid nitrogen tank for longterm storage. After thawing, cells were immediately washed with growth medium to remove any residual DMSO.

3.2.4.2 Transient transfection of suspension cells

Transient transfection of suspension cell line, RLM11, was done by using Neon electroporation system (Invitrogen). One day before transfection, cells were split 1:10 ratio. 10⁷ cells were washed twice with filter sterilized 1X PBS. After removal of supernatant, 10 μ g DNA was added onto pellet and cells were resuspended in 100 μ L of HBS. Mixture was taken into 100 μ L Neon golden tips and placed in electroporation cuvette. Optimum transfection condition for delivery of DNA into cells is 1500 V with a single pulse in 20 milliseconds. Then, cells were transferred to tissue culture flasks containing pre-warmed RPMI supplemented with 10% FBS, 2 mM L-Glutamine, 100 unit/mL, 100 unit/mL streptomycin, non-essential amino acids, vitamin and 50 μ M 2-mercaptoethanol. TALEN transfected RLM11 cells were incubated at 32°C-incubator supplied with 5% CO₂ for 72 hours and harvested for further TALEN genotyping assays.

3.2.4.3 Transient transfection of adherent cells with PEI

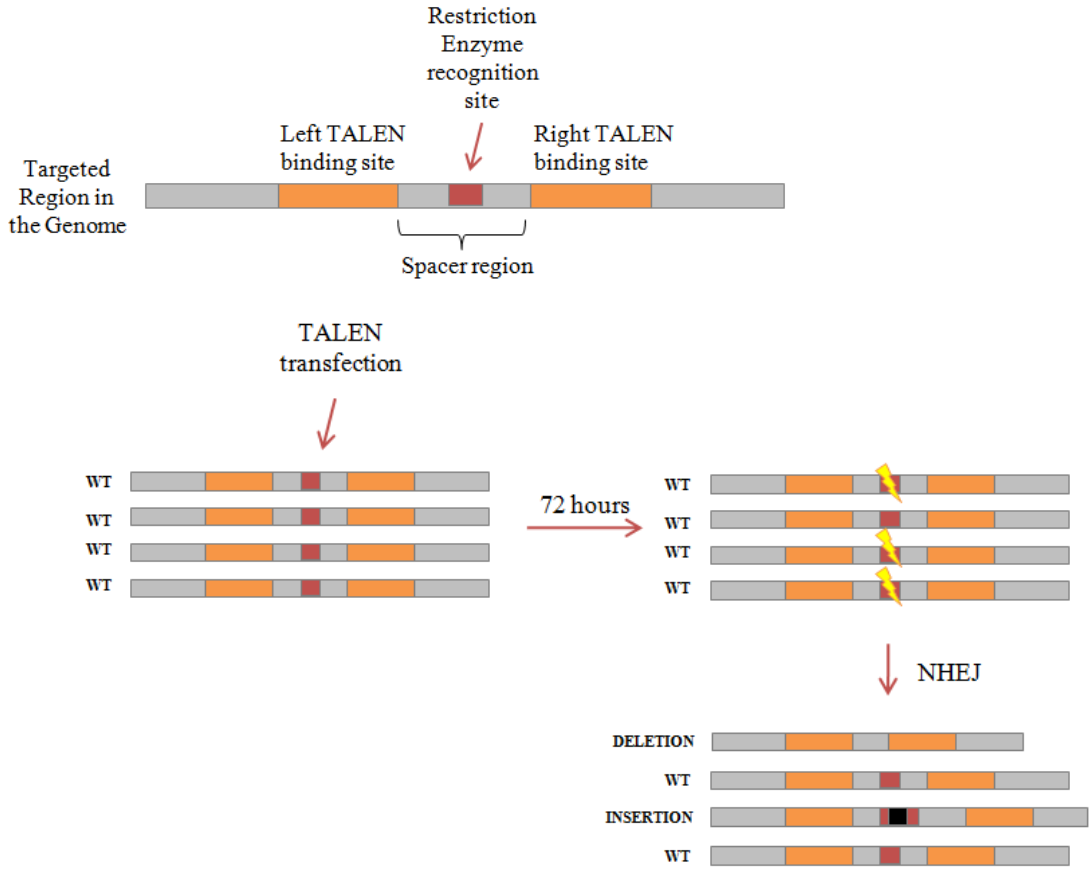
Transient transfection of adherent cell lines was achieved using polyethylenimine (PEI). PEI is a cationic polymer, which forms complex with negatively charged DNA and bind to cell surface. DNA is taken into the cell via endosomal vesicles and osmotic swelling release plasmid DNA to the cytoplasm⁷⁹. One day before transfection, 2.0×10^5 adherent cells (for each well) were split onto 6 well tissue culture plates. On the day of transfection, 3 µg of total DNA was diluted in 200 µL serum-free DMEM without phenol red in a sterile tube. PEI (1µg/µL) was added to diluted DNA based on 3:1 ratio of PEI (µg) to total plasmid DNA (µg) and mixed immediately by vortexing. After 15 minutes of incubation at room temperature, DNA/PEI mixture was added drop by drop on cells in tissue culture dishes. HCT116 cells transfected with TALENs were incubated at 32°C-incubator supplied with 5% CO₂ for 72 hours and harvested for further TALEN genotyping assays.

3.2.4.4 Flow cytometric analysis

10^6 cells were used for each flow cytometric analysis. Flow cytometric analysis of cells was performed using BD FACSCanto. For analysis of cells expressing fluorescent protein GFP, the cells were washed twice with FACS buffer and resuspended in 500µL of FACS buffer for analysis. GFP expression levels were detected with FITC channel. For analysis of IL7R expression, cells were washed twice with FACS buffer and incubated with CD127-Biotin antibody against mouse IL7Rα – Biotin at 4°C for 30 min in dark. After washing twice with FACS buffer to remove unbound antibody, cells were incubated with SA-Alexa647 at 4°C for 30 min in the dark. Cells were washed twice with FACS buffer and resuspended 500µL of FACS buffer for analysis. IL7R expression levels were detected with Alexa-647 channel.

3.2.5 TALEN Induced Mutation Screening

General strategy for detection of mutation at TALEN target site is given in Figure 3.6. In this study, mutation in TALEN target site was detected by the loss of restriction enzyme cut site, via restriction fragment length polymorphism (RFLP) assay.



Genomic DNA isolation of the TALEN transfected pool after 72 hours and PCR amplification of the targeted site; followed by restriction digestion control:

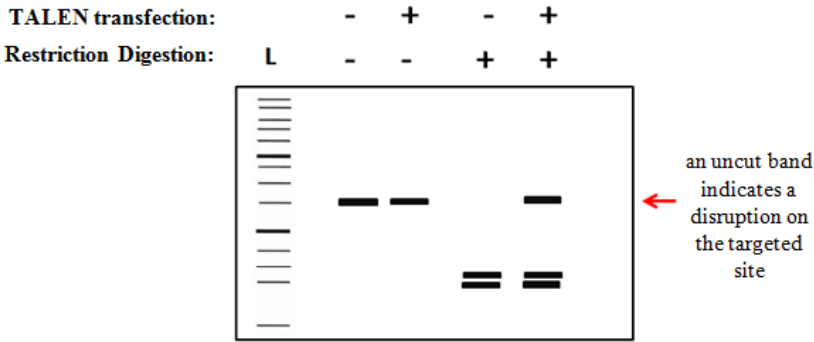


Figure 3.6 General strategy for detection of TALEN induced mutation at target the site.

3.2.5.1 Genomic DNA extraction

72 hours after transfection with TALENs, genomic DNA of the cells was isolated by using GenElute Mammalian Genomic DNA Miniprep Kit (SIGMA).

3.2.5.2 Restriction Fragment Length Polymorphism (RFLP) analysis

Isolated genomic DNA was amplified via PCR reaction before performing digestion. (Optimized PCR conditions given in Table 3.12 using primers for TALEN target site amplification.)

Component	Amount
Template genomic DNA	1.0 μ L
Taq Polymerase Buffer 10X	2.5 μ L
10mM dNTP each	0.4 μ L
Forward primer (10mM)	1.0 μ L
Reverse primer (10mM)	1.0 μ L
Taq Polymerase (5U/ μ l)	0.2 μ L
dH ₂ O	18.9 μ L
Total	25 μ L

Table 3.12 Optimized PCR conditions for TALEN target site amplification

PCR was performed according to following cycle for the respective region:

GR binding site (also NF κ B site):

95°C/4 min + 30X (95°C/30 s + 54°C/30 s + 72°C/60 s) + 72°C/10 min

Notch binding site;

95°C/4 min + 30X (95°C/30 s + 64°C/30 s + 72°C/60 s) + 72°C/10 min

Exon 2 and Exon 3;

95°C/4 min + 30X (95°C/30 s + 60°C/30 s + 72°C/60 s) + 72°C/10 min

hGR TALEN binding site:

95°C/4 min + 30X (95°C/30 s + 62°C/30 s + 72°C/60 s) + 72°C/10 min

PCR products were digested with the enzyme in spacer region and run on agarose gel to detect whether any undigested band is left indicating presence of mutation at TALEN target site.

3.2.5.3 Single Cell Analysis

TALEN expression in mammalian cells induces random and different mutations and most of the cells remain intact after transfection. For that reason single cell analysis was done to amplify one type of mutation and observe its effects on a batch of cells that has the same specific change in their sequence. While growing single cell colonies there are two strategies that can be followed. In the first one after the cells were counted serial dilutions were done in their respective growth medium and when the total number of cells that were desired reached (less than 90 cells for one plate), the medium was divided to 96-well plates. After 7-10 days single cells had enough confluency to transfer to a larger volume. In the second one, certain amount of cells were put into the first lane of the 96-well plate and serial dilutions were done in the 96-well plate itself. Only the cells that were grown right before the empty wells were chosen to do further experiments. Figure 3.7 explains the dilution methods.

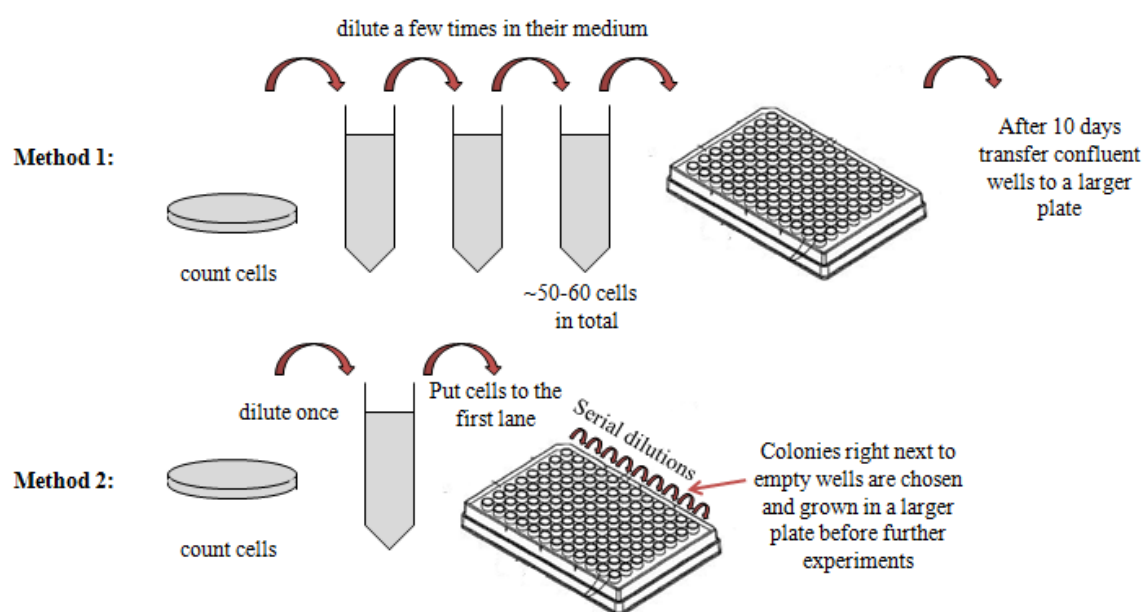


Figure 3.7: Methods for obtaining single cell colonies.

After growing single cell colonies mutants are detected either by RFLP analysis after genomic DNA isolation or by FACS analysis by comparing their IL7R expression levels to the WT (only for the experiments done with RLM11 cell line).

In order to sequence suspected mutants gel extraction was performed for the band in size of the undigested PCR product (in case there is still a band that can be digested) and cloned using CloneJETTM PCR Cloning Kit (Thermo Scientific) or InsTA Clone PCR Cloning Kit (Thermo Scientific). Uncut bands were cloned to pJET1.2/blunt or pTZ57R/T vector and colony PCR was performed to ensure presence of insert in selected colonies according to conditions provided by the kit. Plasmid DNA was isolated from 3-ml overnight cultures of true colonies and sequenced to evaluate mutations generated at the cleavage site.

RESULTS

4.1.Targeting IL-7R Gene

4.1.1. Targeting Transcription Factor Binding Sites of the IL-7R Gene

In this study we used TALEN genome editing tools to mutate two of the transcription factor binding sites in the IL-7R enhancer region, the glucocorticoid receptor (GR) binding site and the Notch binding site (Figure 4.1). First, we constructed two pairs of different TALEN proteins targeting the GR binding site of the IL-7R gene using the Golden Gate TALEN Assembly procedure as explained in the Methods section. The TALEN pair targeting the Notch binding site was designed by Şeyda Temiz ⁷⁸. Using the same method we expressed both Notch TALENs and GR TALENs in the IL-7R positive murine cell line RLM11 to observe the effects of mutations occurring in these regions on the expression of the IL7R gene. We hypothesized that if the binding of one of these transcription factors to their corresponding sites had a positive critical role in the expression of the IL-7R protein, the introduced mutations would decrease the IL-7R expression levels in these cells.

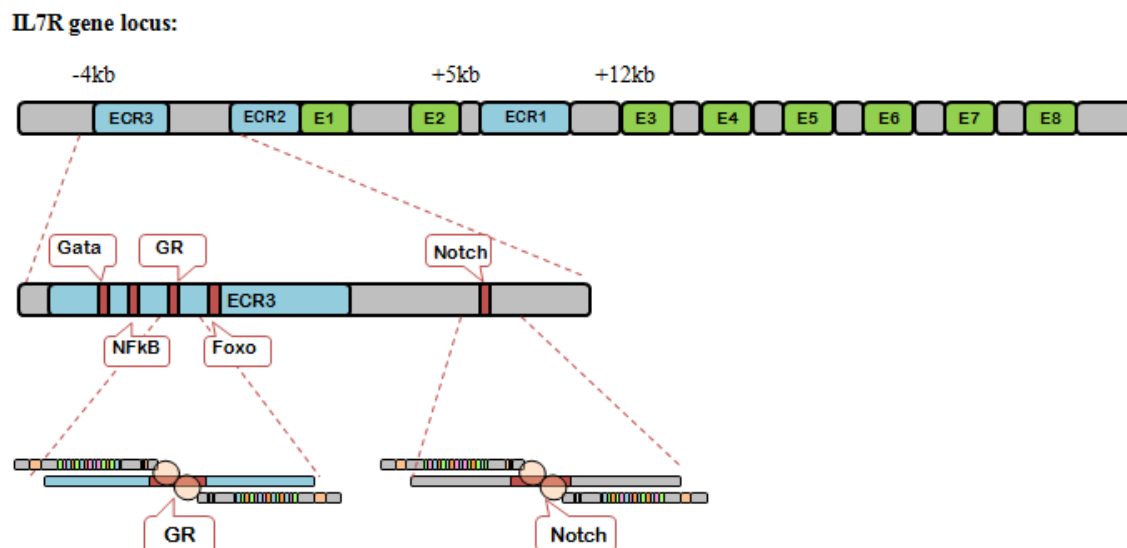


Figure 4.1 Schematic representation of the mouse IL7R α gene locus. The IL7R α gene contains eight exons and three evolutionarily conserved regions (ECR). ECR3 has binding sites for Gata, NFkB, GR and Foxo. Notch binds upstream of the ECR2. Two transcription factor binding sites targeted in this study were the GR and Notch binding sites.

4.1.1.1. TALENs Targeting GR Binding Site of IL-7R

We designed and constructed two pairs of TALENs for this experiment; one of the pairs, named as GR TALEN2 directly targeting the GR binding site, and the other pair, GR TALEN3 targeting a slightly shifted region, placing an MboII restriction enzyme recognition site in the middle. Since we use the restriction fragment length polymorphism (RFLP) assay to detect mutations as summarized in Figure 3.6, we assumed selecting mutated colonies would be easier in case the TALEN3 pair was used, while the TALEN2 pair would be more successful in mutating the GR binding site. Their binding sites and the targeted region are shown in Figure 4.2.

While designing the TALEN pairs, we paid attention to having 15 to 16 bp long spacers in between the pairs since it has been shown that longer or shorter spacer length results in reduced functional activity. In addition, having a 5'T base preceding the TALEN pair binding sites is important for their binding efficiency; therefore, all of the TALENs we designed have a T nucleotide right before the first base they bind. Since the right TALENs, or reverse TALENs, bind the complementary strand, in Figure 4.2

the preceding base appears to be adenine. Both monomers of the TALEN2 pair binds a region that is 15 nucleotides long while the TALEN3 pair binds a region that is 16 nucleotides long.

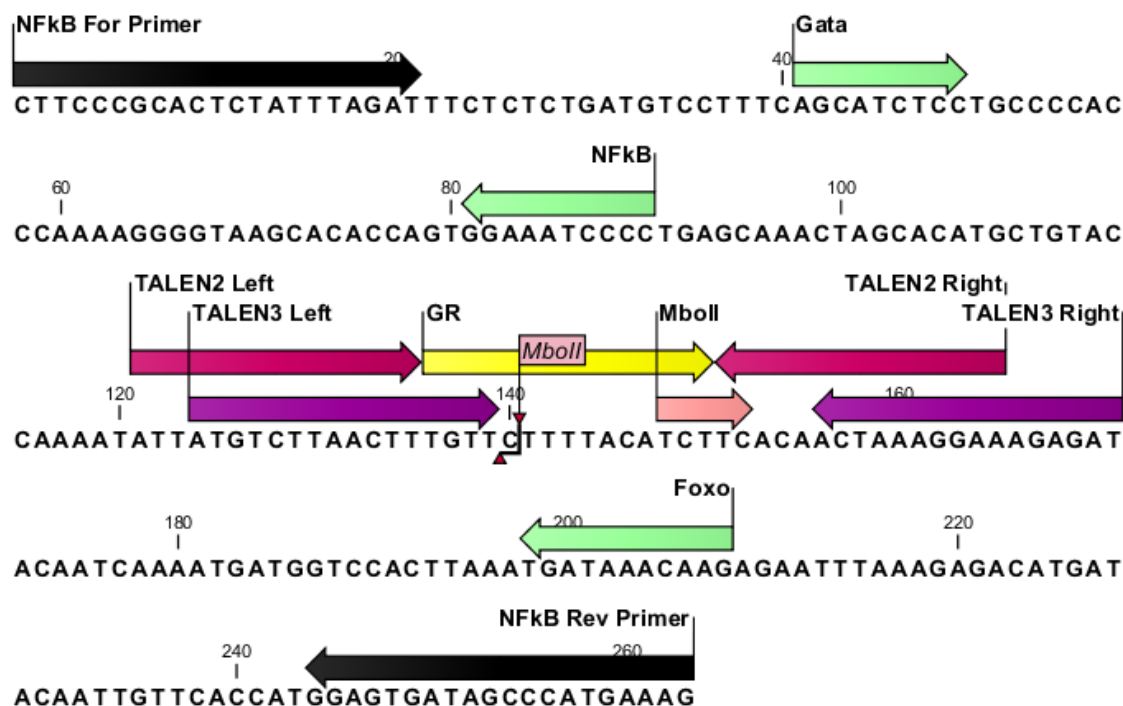


Figure 4.2 Binding sites for the TALENs targeting the GR binding site. The yellow arrow indicates the GR transcription factor binding site, pink and purple ones indicate the left and right TALEN pairs that were numbered as 2 and 3. Light pink arrow shows an MboII restriction enzyme recognition site, and it cuts a site 7 nucleotides away. Green arrows indicate the neighboring transcription factor binding sites and the black arrows show the primers used to amplify this region.

4.1.1.1.1. Assembly of GR Site Targeting TALENs

In order to construct the TALENs targeting the GR binding site, we used the Golden Gate TALEN Assembly kit as described in the Methods section (Figure 3.4). TALEN assembly takes 5 days and involves two main reactions for the assembly of individual repeat modules into array plasmids and assembly of array plasmids into a final mammalian expression plasmid. We used NI, HD and NG RVD modules to recognize and bind A, C and T nucleotides respectively and as G nucleotide recognizing RVD we used NH, differing from the previously designed TALENs in our lab (as was

the case for the Notch site RVDs). This modification was because of recent experiments that showed that NH RVDs are more specific in binding G than NN¹⁵.

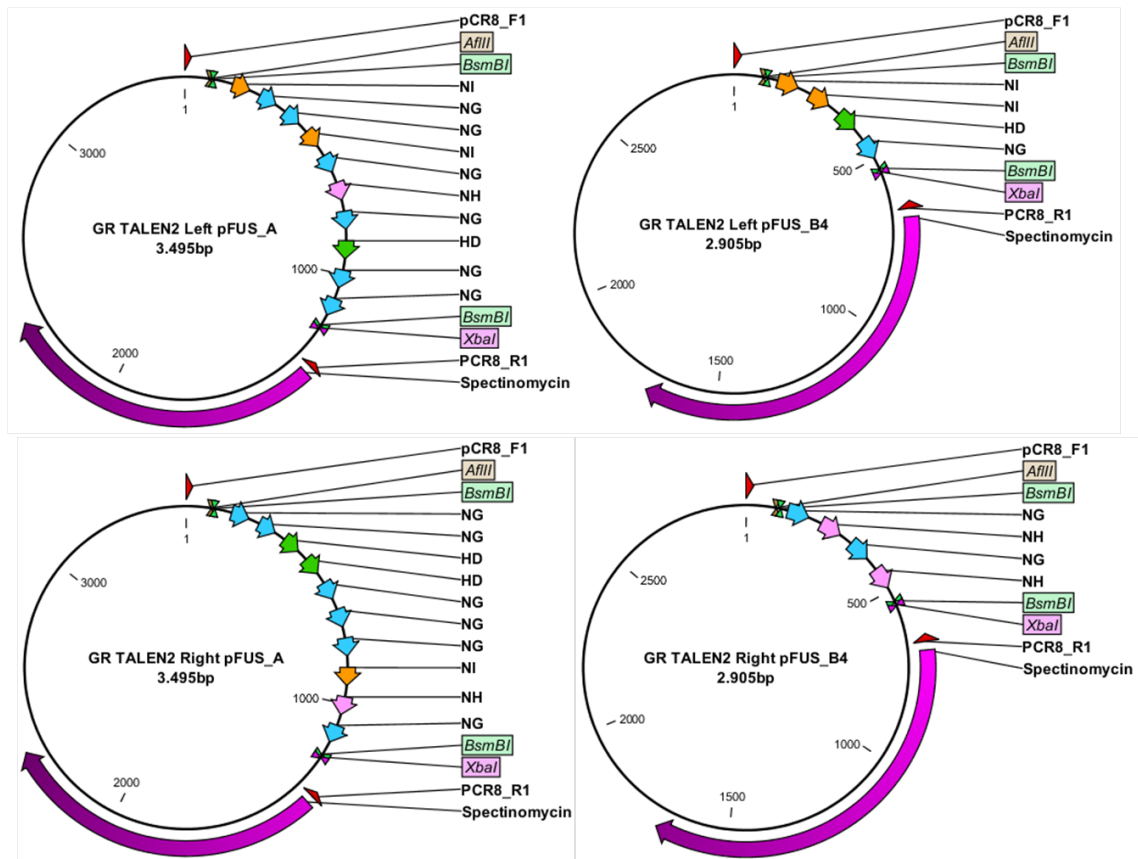


Figure 4.3 Plasmid maps showing GR TALEN2 Left and Right TALEN pair repeats in pFUS_A and pFUS_B4 intermediary plasmids after Golden Gate Reaction #1.

For the construction of four different TALENs with 15 and 16 repeats we cloned first 10 repeats in pFUS_A plasmids and the remaining ones except the last ones were cloned into pFUS_B4 and pFUS_B5 (Golden Gate Reaction #1). Therefore, we set up two Golden Gate reactions (reaction A and B) for each monomer of the TALEN pair. We transformed the Golden Gate reaction #1 into *E. coli*, plated the reactions on agar plates with spectinomycin and used IPTG/X-gal for blue-white screening. We performed colony PCR on white colonies using the primers pCR8_F1 and pCR8_R1. For repeats in the pFUS_A plasmids, we detected bands around 1200 bp, whereas for the pFUS_B4 and pFUS_B5 plasmids, we obtained bands around 600 bp and 700 bp respectively. In addition to the band of expected size, a smear and ladder of bands were also detected which results from the presence of repeats in clones. After isolation of plasmid DNA from two different correct colonies, we performed an AflIII-XbaI

diagnostic digest which should generate bands reflecting the number of repeats such that a 1048bp-band was obtained for a pFUS_A plasmid with 10 repeats. Similarly the band sizes were 430bp and 523 bp for the pFUS_B4 and pFUS_B5 plasmids with four and five repeats. Figure 4.3 and Figure 4.4 show plasmid maps related to reactions A and B of the GR TALEN2 and GR TALEN3 left and right pairs. Figure 4.5 shows agarose gel images after colony PCR with pCR8_F1 and pCR8_R1 primers and confirmation of those colonies with AflIII-XbaI double digests.

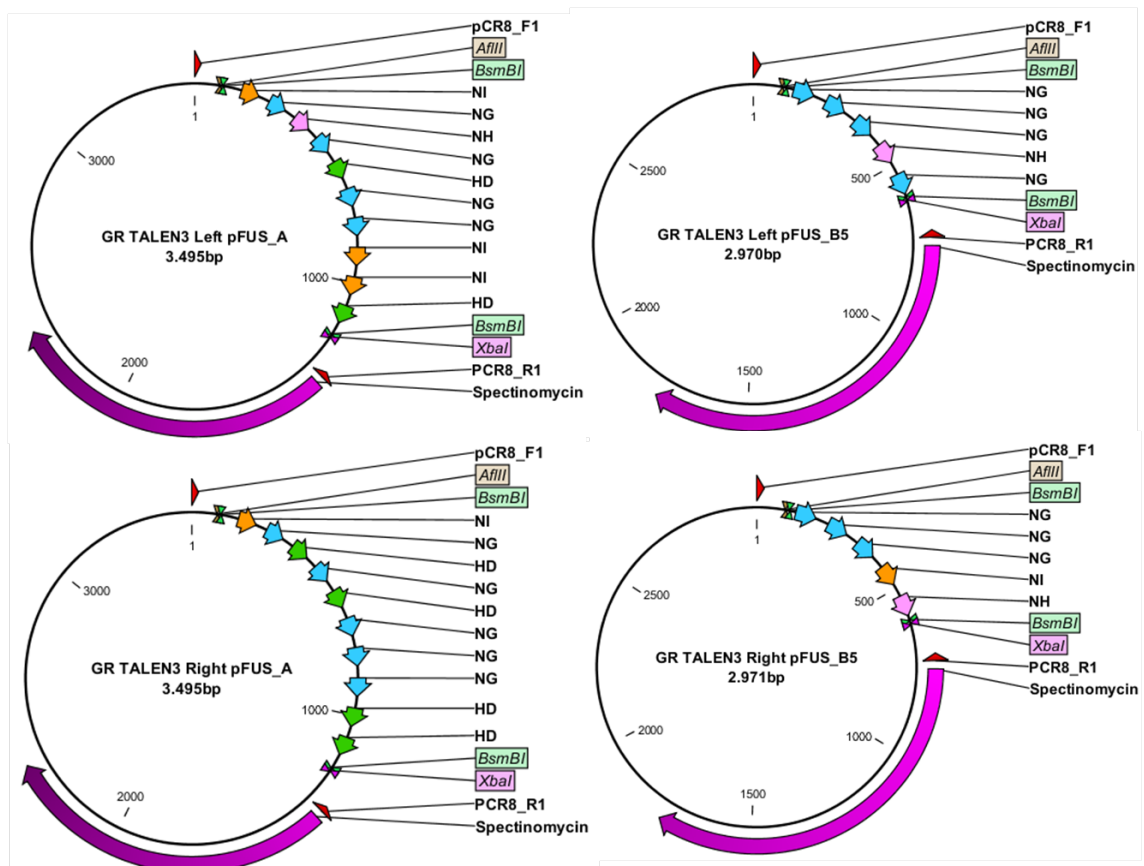


Figure 4.4 Plasmid maps showing GR TALEN3 Left and Right TALEN pair repeats in pFUS_A and pFUS_B4 intermediary plasmids after Golden Gate Reaction #1.

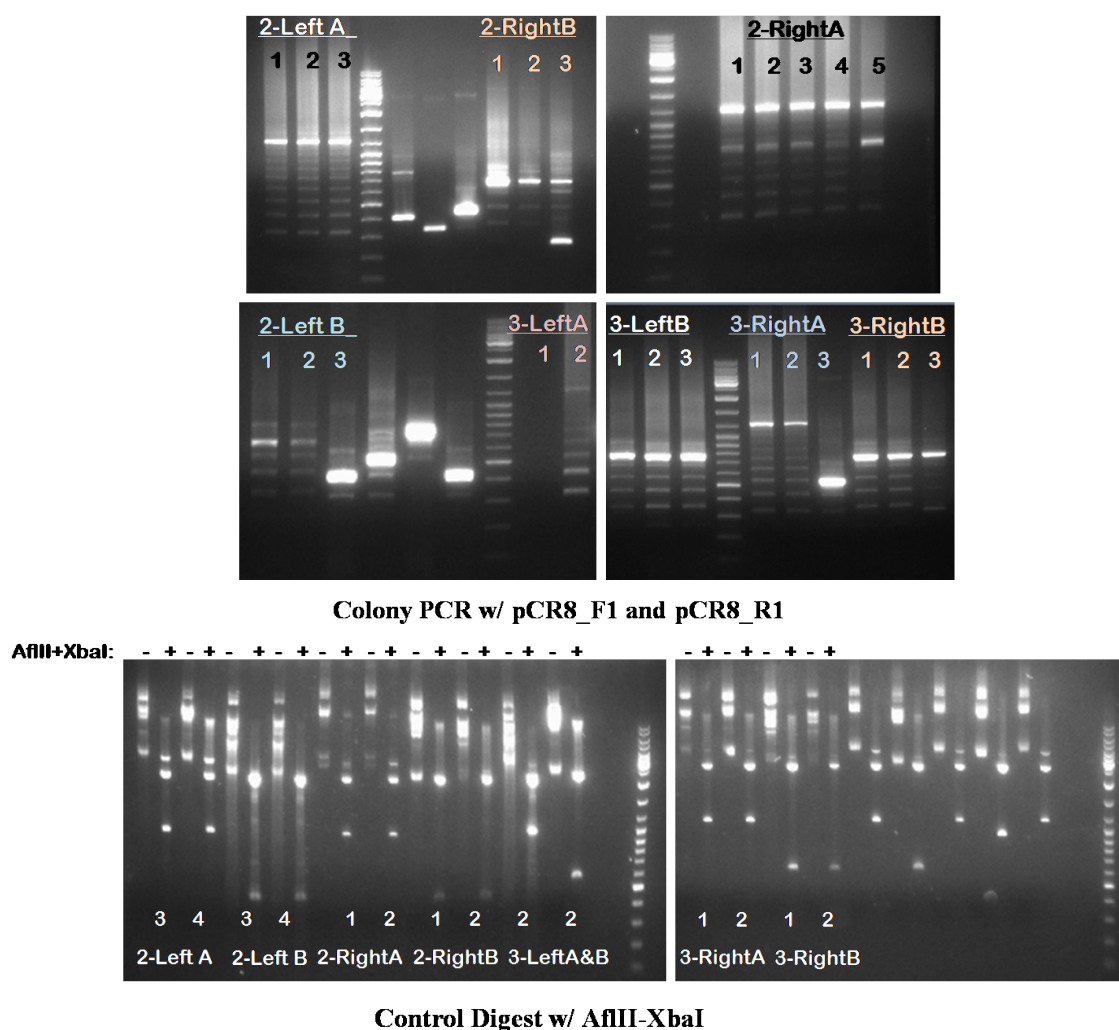


Figure 4.5 Agarose gel images showing colony PCR results and control digests done to find correctly assembled array plasmids for both TALEN2 and TALEN3 pairs after Golden Gate Reaction #1.

After the construction of intermediary array of repeats were complete, we joined repeats coming from reactions A and B, the last repeat plasmids and the final backbone plasmid pC-GoldyTALEN in Golden Gate reaction #2. This time after transformation the colonies were grown on ampicillin agar plates and we performed colony PCR using primers TAL_F1 and TAL_R2, which gives bands around 1740bp for plasmids with 15 repeats and 1840bp bands for plasmids with 16 repeats. Along with the correct bands a ladder effect and a smear is expected. Consistent with this, we found that the colonies that give a single band were not correctly assembled. After isolating plasmid DNAs of correct colonies we performed a confirmation digest with the AatII and StuI enzymes,

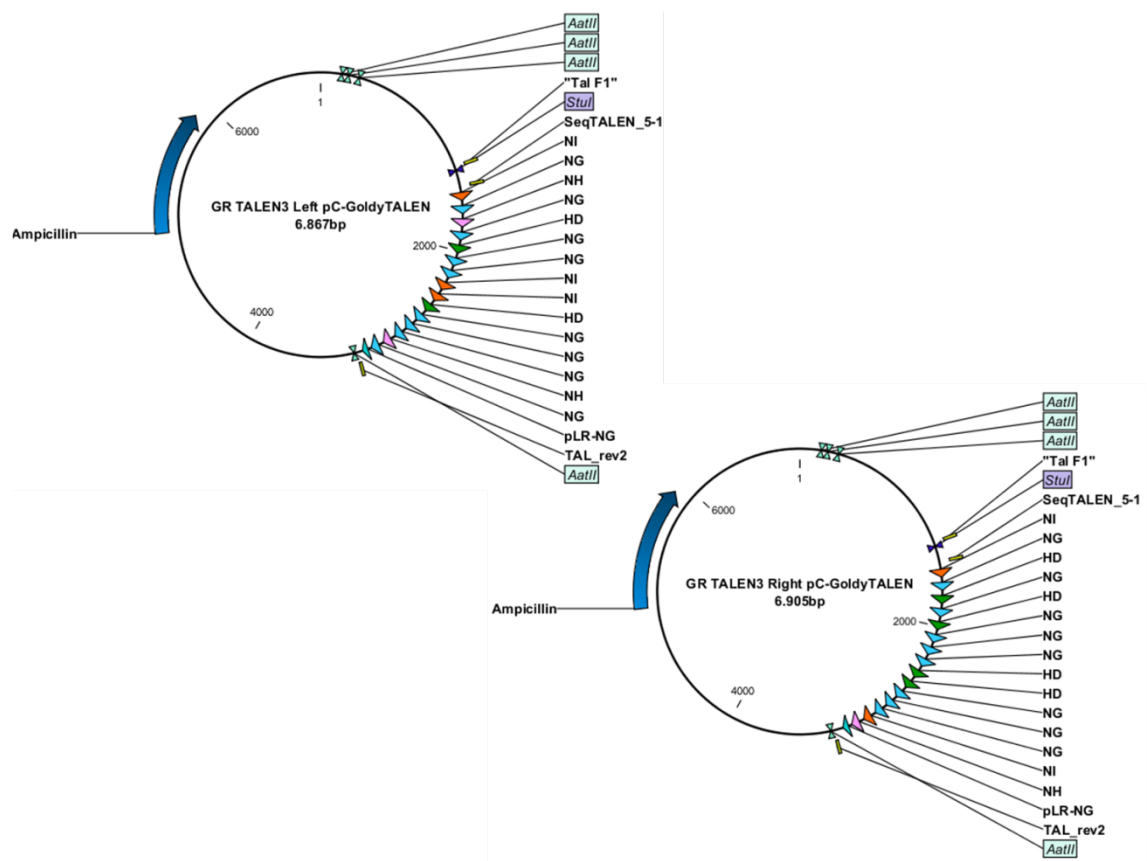


Figure 4.7 Plasmid maps showing fully assembled GR TALEN3 Left and Right TALEN pair in pC-GoldyTALEN backbone.

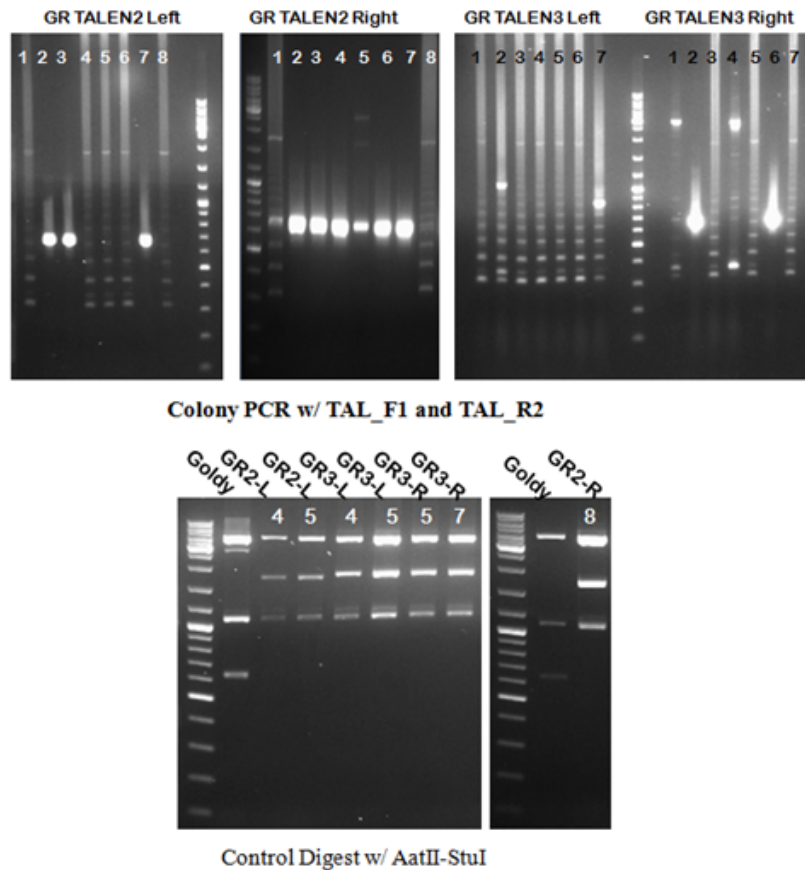


Figure 4.8 Agarose gel images showing colony PCR and control digest results of GR TALEN2 and GR TALEN3 pairs after Golden Gate reaction #2.

4.1.1.1.2. Expression of the Constructed GR TALEN Pair in RLM11 cells and Detection of Site Specific Mutations

After we completed the construction of two different TALEN pairs targeting the GR binding site of the IL-7R gene, we wanted to confirm their functionality. We used the RLM11 cell line, a mouse thymoma cell line which has a high level of IL-7R expression on its surface. We assumed that any damage induced on the gene control elements would easily be observed on the expression phenotype of the cell surface protein. We first transfected RLM11 cells with the GR TALEN3 pair using the Neon electroporation system. Because this TALEN target site had an MboII restriction enzyme binding site in its spacer region, we expected to find mutations induced easier than the second TALEN pair whose activity cannot be assessed using RFLP. After the cells were transfected, we incubated the cells in 32°C for 72 hours and at the end of this period we analyzed the cells by flow cytometry (Figure 4.9). We have observed that

even though IL-7R levels dropped for GR TALEN transfected cells, the cells we transfected with pcDNA-GFP plasmid for transfection efficiency control had lower IL7R levels as well. We observed this result probably because the cells could not regain their health after electroporation and 32°C incubation. Therefore we could not relate this result to TALEN function.

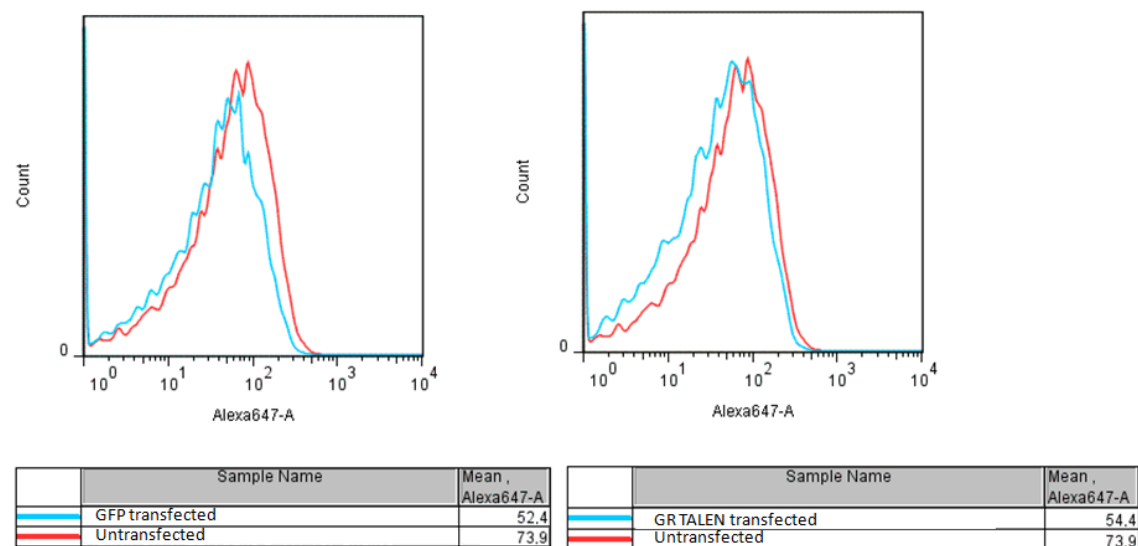


Figure 4.9 IL-7R expression levels of pcDNA-GFP and GR TALEN3 transfected RLM11 cells compared to untransfected cells. Histogram for Alexa-647 represents surface IL-7R α expression.

Since IL-7R levels only could not indicate the existence of a mutation, we extracted the genomes of both untransfected and TALEN transfected samples and performed the RFLP assay (as explained in Figure 3.6). We PCR amplified the region with the indicated primers in Figure 4.2 and digested the amplicons with the MboII restriction enzyme. We have observed an uncut band in TALEN transfected samples, indicating that a large portion of the cells were mutated (Figure 4.10).

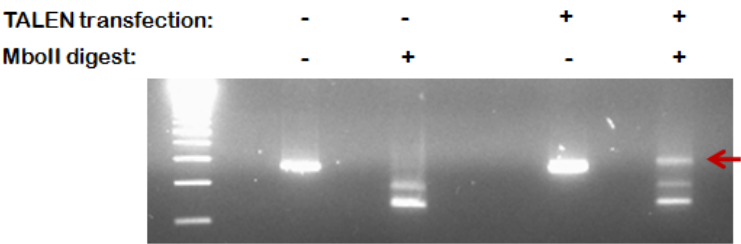


Figure 4.10 RFLP analysis on the GR TALEN transfected and untransfected RLM11 cells. The uncut band the red arrow indicates the presence of mutation.

4.1.1.1.3. Single Cell Screening of GR TALEN transfected RLM11 cells and Detection of Mutants

Since we could not detect any significant changes in the phenotype of the TALEN transfected pool of cells, and assuming that the presence of unmutated cells in the pool could be preventing the detection of mutated cells, we decided to grow single cell colonies from this pool and screen for cells that have mutations, preferably on both chromosomes. The strategy to obtain single cell colonies is explained in Figure 3.7. After a portion of the single cells we plated grow, we analyzed their mutation status by RFLP and separated the ones that had uncut bands (Figure 4.11).

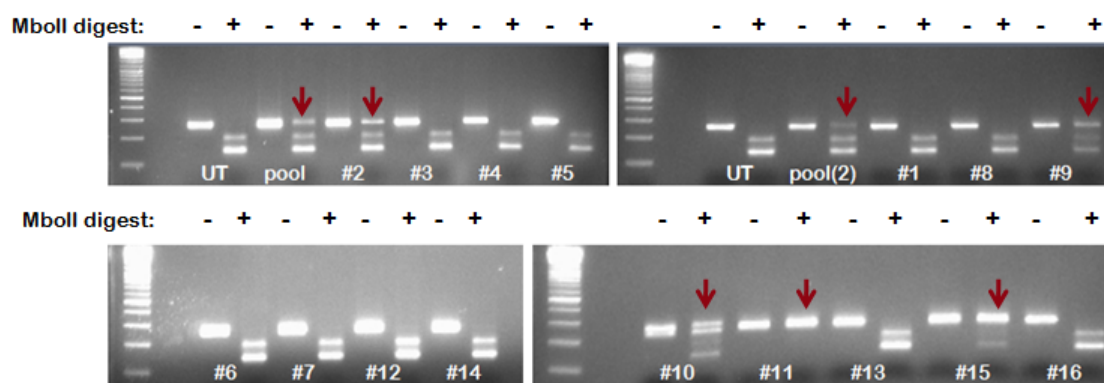


Figure 4.11 RFLP results for GR TALEN transfected single cell colonies. Red arrows show the uncut bands caused by the presence of a mutation in the restriction enzyme recognition site. UT stands for untransfected.

After repeating the experiment to ensure that all of the colonies we selected are mutants, (Figure 4.12) we gel extracted the uncut bands and cloned them into the pTZ57R/T plasmid using the InsTAclone PCR cloning kit to send them to MCLAB Inc. (San Francisco, CA, USA) for sequencing.

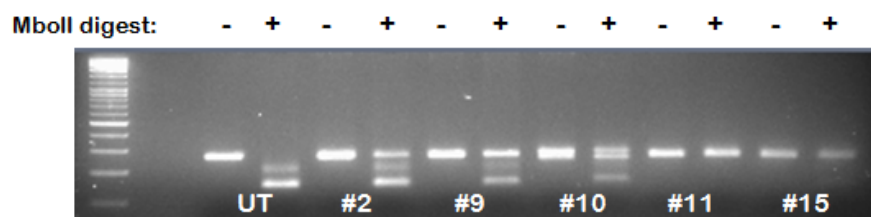


Figure 4.12 Selected RLM11 single cell colonies to send sequencing.

Sequencing results of the selected mutant colonies demonstrated the presence of several types of mutations (Figure 4.13). All of the mutations had disrupted the GR binding site, however, the 2nd, 9th and 15th colonies still had an intact wild type sequence along with a mutated allele.

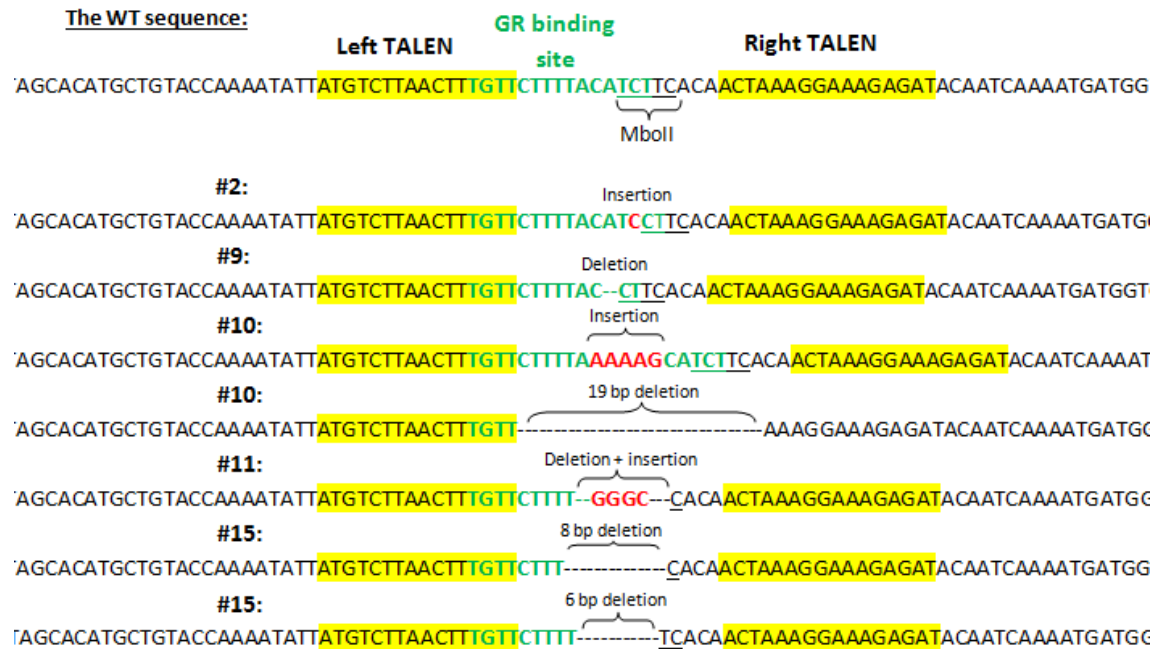


Figure 4.13 Sequencing results of single cell colonies that had uncut bands in the RFLP assay. Yellow highlighted sequences indicate TALEN binding sites, green letters indicate the GR binding region, underlined sequence is the MboII restriction enzyme binding site and red letters show the insertion sequence.

4.1.1.1.4. IL-7R Expression in GR binding site Mutant RLM11 cells

After selecting the mutant single cell colonies that had their GR binding site disrupted at least in one of their chromosomes, we wanted to compare their IL-7R expression levels to wild type RLM11 cells. While one of the single cell colonies, the 10th colony, had shown lower levels of IL7R expression compared to wild type consistently, the rest of the colonies had IL7R expression around the same levels as the wild type, and some of them had even significantly higher expression levels in repeated experiments. An example FACS analysis from each colony is shown in Figure 4.14.

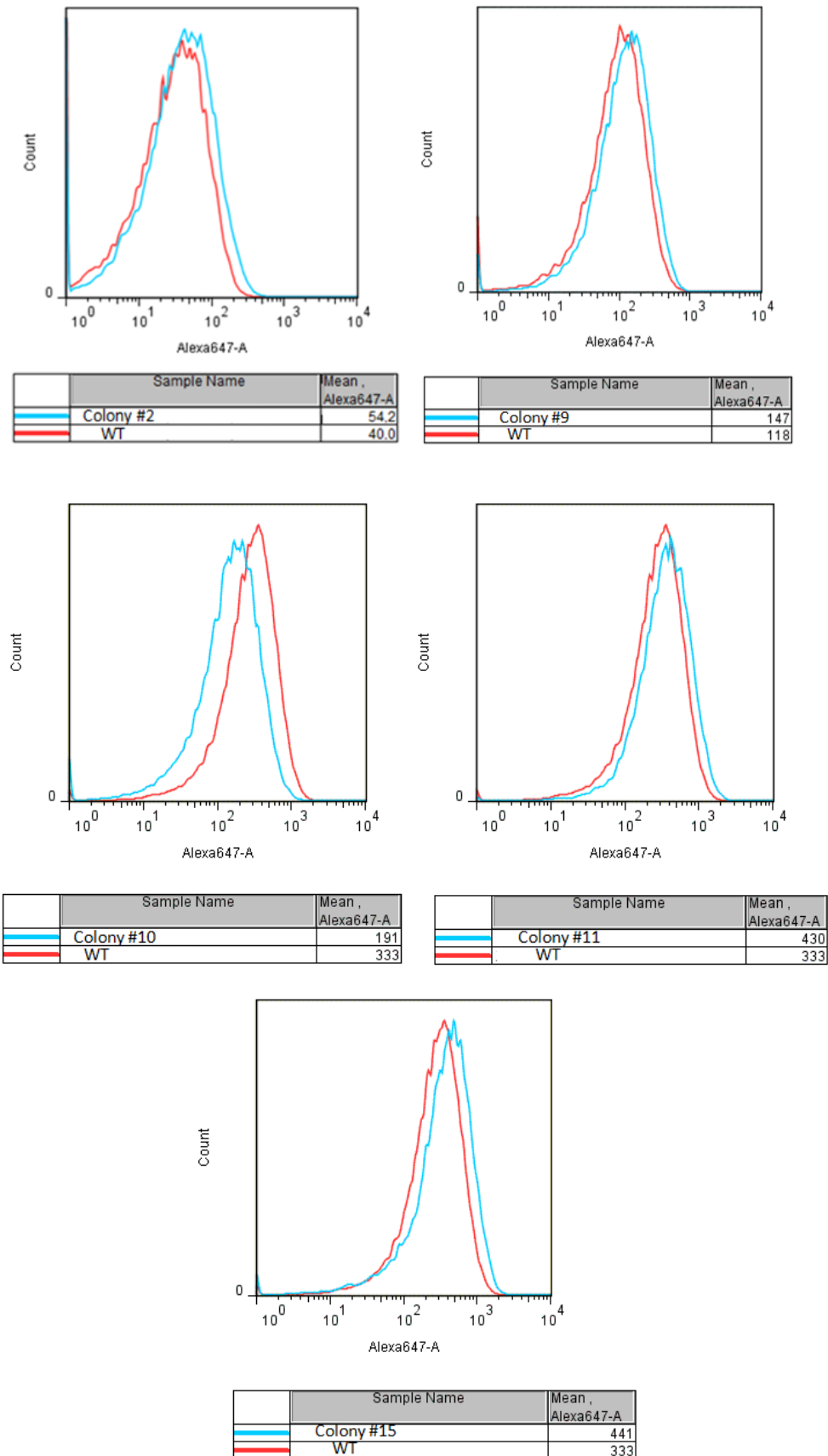


Figure 4.14 FACS analysis showing IL7R expressions of mutant RLM11 single cell colonies. Histogram for Alexa-647 represents surface IL-7R α expression.

4.1.1.2 TALENs Targeting Notch Binding Site of IL-7R gene

In a previous study done in our lab IL7R Notch binding site targeting TALENs were assembled and different mutations caused by imprecise repair of double strand breaks induced by these TALENs were analyzed in Neuro-2a and RLM11 cell lines⁷⁸. The experiments were performed in a pool of cells that contained a majority of wild type cells along with the cells with various mutations induced by TALEN transfection, which made them unsuitable for IL7R expression experiments. In this study, we wanted to continue this project by growing single cell colonies from these TALEN transfected pools and observe the effect of specific Notch site mutations on IL7R expression levels. The Notch TALEN pair binding sites and the targeted region is shown in Figure 4.15.

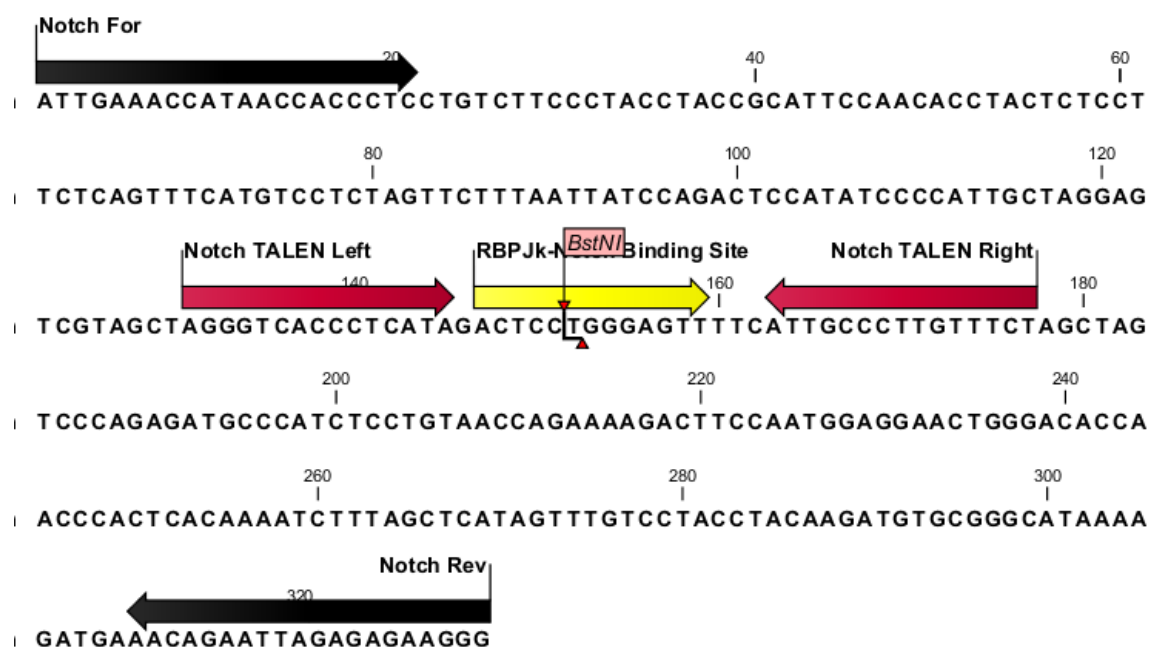


Figure 4.15 The binding sites for IL-7R Notch site targeting TALEN pair. Yellow arrow indicates the RBPJk-Notch binding site, the flag in between shows the sequence restriction enzyme BstNI cuts, the black arrows indicate the primers that were used to amplify this region.

4.1.1.2.1 Expression of a previously designed Notch TALEN pair in RLM11 cells and mutation screening via expression of IL-7R in single cell colonies

We transfected the RLM11 cell line with the expression constructs encoding Notch TALEN pair and compared the IL-7R expression levels of untransfected and transfected cells by FACS analysis and observed that even though Notch TALEN transfected cells had lower expression, it was not significant (Figure 4.16).

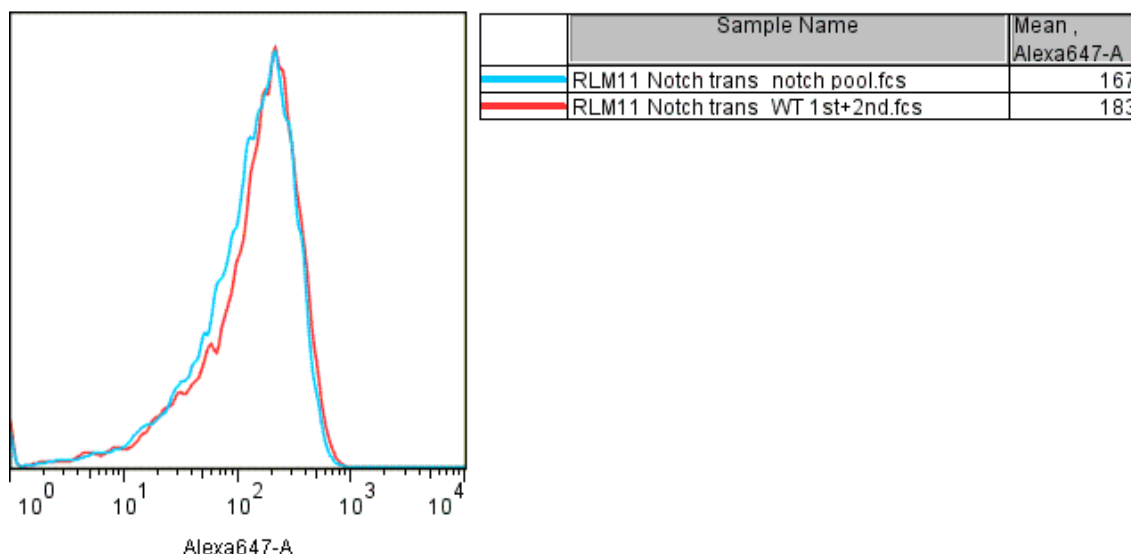


Figure 4.16 IL-7R expression levels of Notch TALEN transfected RLM11 cells compared to untransfected cells. Histogram for Alexa-647 represents surface IL-7R α expression. Red histogram shows WT expression while blue histogram shows the IL7R expression on transfected cells.

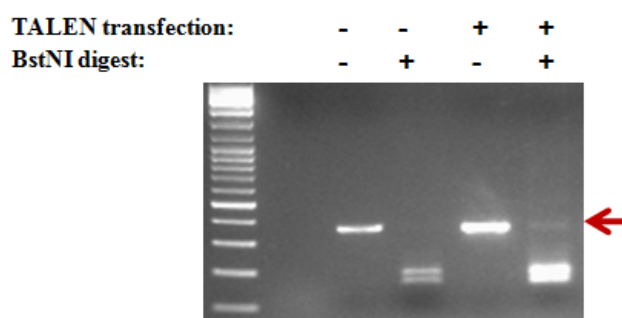


Figure 4.17 RFLP analysis on the Notch TALEN transfected and untransfected RLM11 cells. The uncut band which is pointed by a red arrow indicates the presence of a targeted mutation.

Along with FACS analysis we checked the presence of a mutation by using the RFLP assay (Figure 4.17) and even though we observed an uncut band the percentage

was too low compared to our GR TALEN transfected cells. For that reason while screening for mutant colonies among single cells instead of extracting genomic DNA from hundreds of colonies and performing the RFLP assay we followed a different approach. We expected that the Notch binding site mutated cells would have lower expression of the IL-7R protein if Notch had a positive role in the expression of IL7R. Therefore, we screened only colonies that had lower IL7R expression levels. We predicted that this approach would increase our chance of finding mutated colonies. Single cell colonies were grown as explained in Figure 3.7 using the first method and after about two weeks we performed FACS analysis on 91 single cell colonies. We selected colonies that had significantly lower IL-7R expression to continue with further experiments (Figure 4.18).

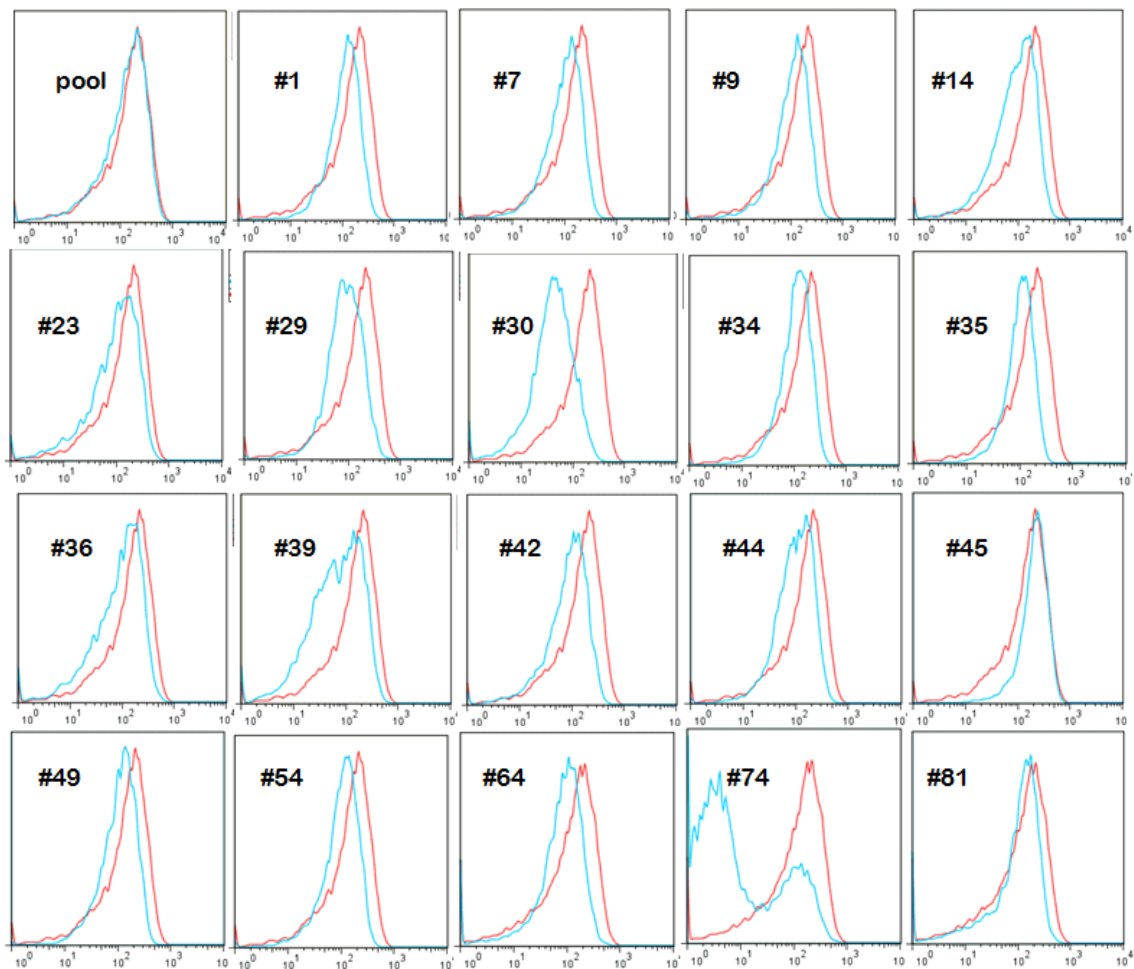


Figure 4.18 IL-7R expression levels of Notch TALEN transfected RLM11 single cell colonies that have lower expression levels compared to WT expression. Histogram for Alexa-647 represents surface IL-7R α expression. Red histograms show the expression on untransfected cells while blue histograms show the expression levels on colonies being examined.

4.1.1.1.2. Detection of mutation in suspected Notch TALEN transfected single cell colonies

After eliminating a large portion of the candidate mutant single cell colonies due to their existing IL-7R expression we predicted that it would be easier to find mutated single colonies by performing the RFLP assay. We extracted genomic DNA from these colonies and digested with the BstNI enzyme which had a cut site in the middle of Notch binding sequence. According to RFLP, from around 20 colonies only two appeared to be mutated; 35th and 49th (Figure 4.19).

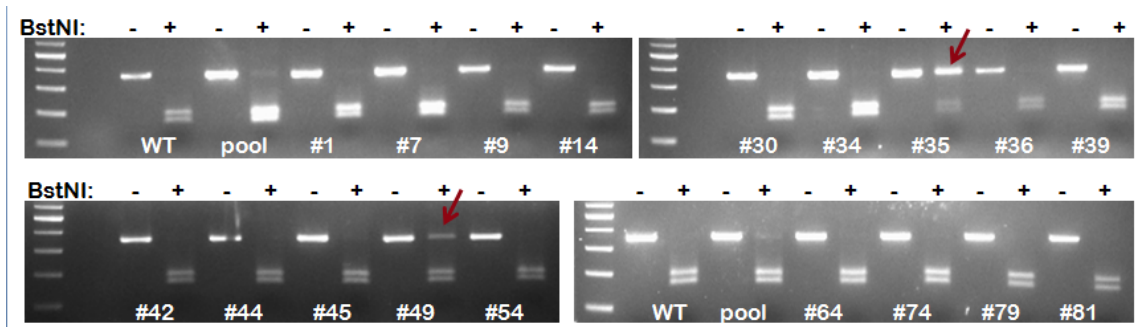


Figure 4.19 RFLP assay for Notch TALEN transfected RLM11 single cell colonies. The restriction digestion was done with the BstNI enzyme and the red arrows show the uncut bands which indicate the presence of mutations.

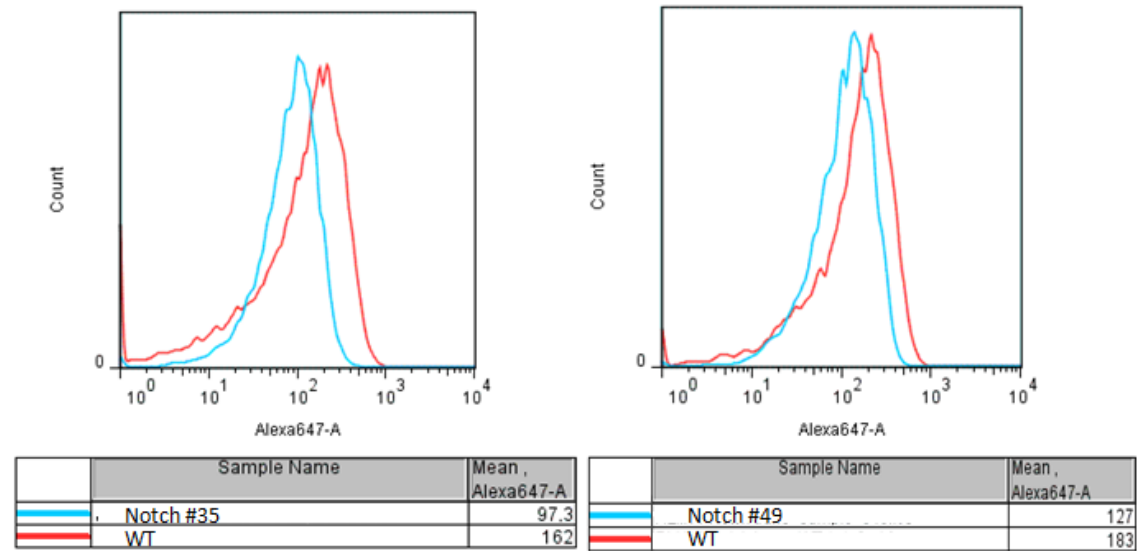


Figure 4.20 IL-7R expression levels for Notch binding site mutated RLM11 cells. Histograms for Alexa-647 represents surface IL-7R α expression.

The single cell colonies that were confirmed to be mutated through RFLP analysis had significantly lower IL-7R expression compared to IL-7R expression on WT RLM11 cells (Figure 4.20). However, the single colonies #30 and #74 showed the lowest expression levels. We examined the type of mutations that disrupted the Notch binding site of the colonies with uncut bands in the RFLP assay. It is possible that some colonies contained mutations that we could not detect by RFLP. Therefore, we cloned PCR products that contained the mutation region into the pJET1.2/blunt plasmid using the CloneJET PCR kit and sent 8 different colonies for sequencing. Sequencing analysis showed that the 35th colony had two types of mutation, a deletion and an insertion; along with an intact WT binding site. The 30th and the 74th colonies appeared to be all WT and 49th colony despite having an uncut band, had only PCR products with the undisrupted binding site cloned (Figure 4.21).

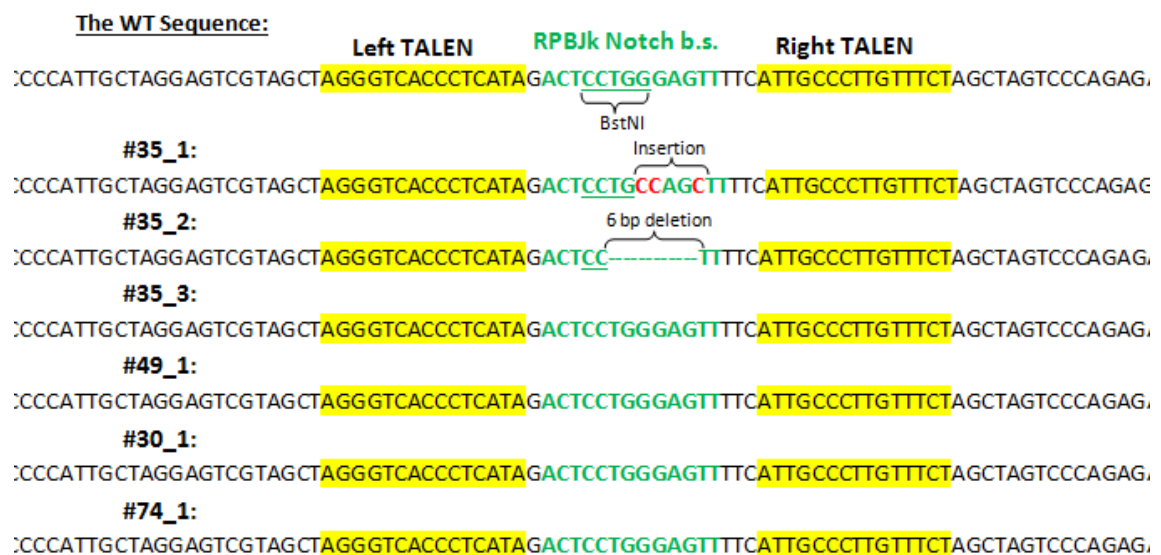


Figure 4.21 The sequencing results for Notch site mutated single cell colonies #35, #49, #30 and #74. The green letters indicate the Notch binding site, yellow highlighted letters show the regions TALEN pair binds, underlined letters are the recognition site for BstNI restriction enzyme.

Among the single cell colonies that had lower IL-7R levels compared to WT expression the 74th colony was the lowest one, and in contrast to others it was the only one that had a complete down-regulation profile. For that reason, we hypothesized that this could be due to a much larger deletion in the genome than we could amplify with our Notch primers. We performed PCR with primers that were used to amplify ECRs (shown in Figure 4.22) along with the Notch primers to see if any of the sites included

one of the Notch primers were deleted, however the 74th colony still gave the same band as the WT (Figure 4.23). If the deletion still exists and both of the Notch primers are deleted, these experiments would not give the results we expected to see.

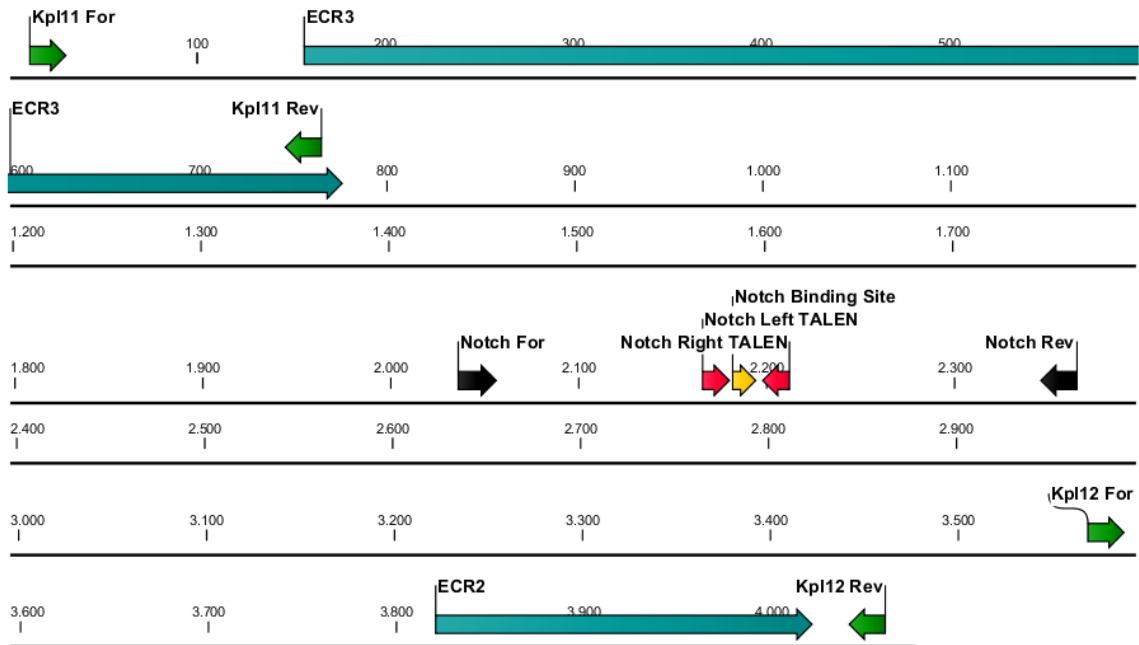


Figure 4.22 The IL7R gene ECR2-ECR3 locus. Blue arrows show ECR2 and ECR3 of the IL7R gene, the Notch binding site is in between these regions indicated by a yellow arrow, red arrows at the left and right sides are TALEN binding sites. Primers that are used to amplify the Notch site are shown with black and primers that are used to amplify ECR2 and ECR3 are shown with green arrows.

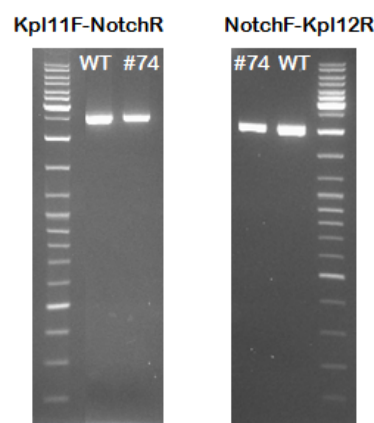


Figure 4.23 PCR amplification of the 74th colony with different primers. Kpl11 forward and Notch reverse primers give bands around 2360bp and Notch forward and Kpl12 reverse primers gives band around 2020bp in WT genome. The binding sites of the primers are shown in Figure 4.22.

4.1.2. Use of TALENs to Delete an Entire Intronic Region of IL-7R gene

One of the evolutionarily conserved regions of the IL-7R gene, ECR1, is in between the second and the third exons. Along with the transcription activator binding sites in the upstream enhancer region of this gene we wanted to target this ECR site also, but differing from the previous experiments instead of targeting one specific binding site, we planned to delete the whole intron in between the 2nd and the 3rd exons using two TALEN pairs simultaneously. We designed the TALEN pairs to target the exons directly, assuming that if both pairs cut the DNA at the same time, the site in between them would be removed. To ensure that the exons remained intact while the intron site is removed, we used a donor DNA which had exon 2 and 3 fused to each other so that when the double strand breaks were induced by the TALEN pairs at the same time, the region would be repaired according to the donor DNA by homologous recombination. The strategy to remove intron 2 of the IL7R gene by TALENs is explained in Figure 4.24.

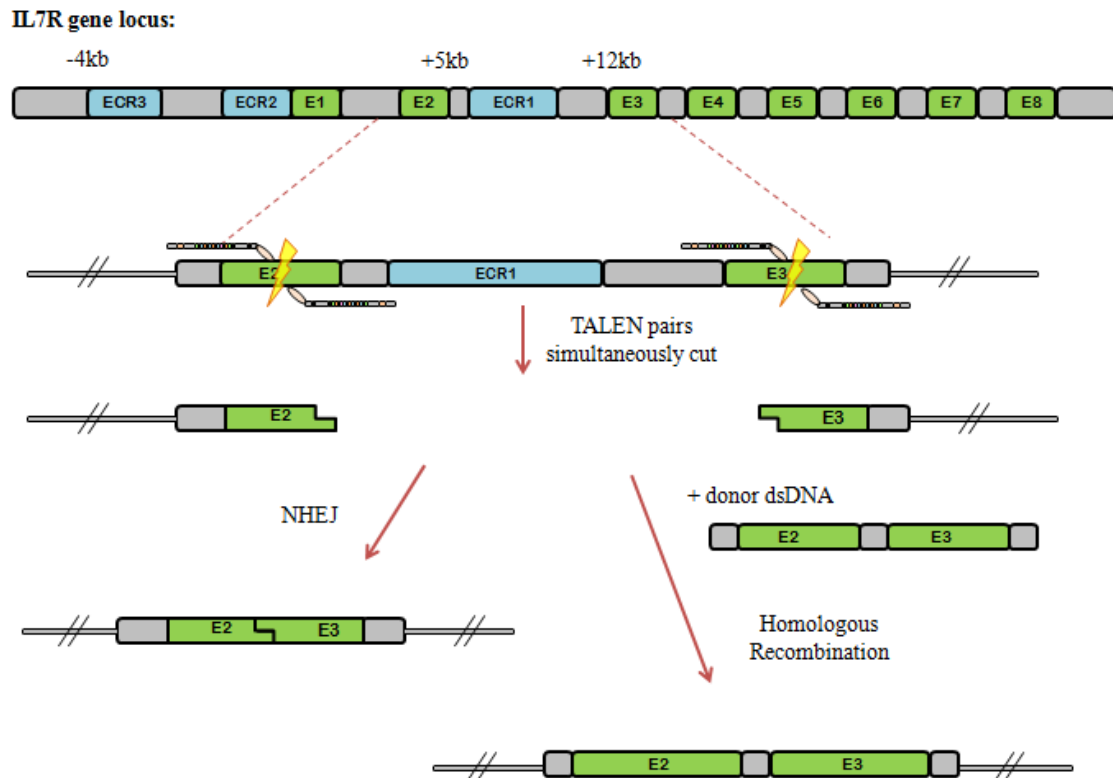


Figure 4.24 The strategy to delete the ECR1 region of the IL-7R gene using two TALEN pairs and a homologous donor DNA.

4.1.2.1. Assembly of TALENs targeting Exon 2 and Exon 3 of the IL-7R gene

In order to remove the region in between exon 2 and 3 we designed two separate TALEN pairs to target random sites in the middle of the exons. As we did with the previous designs we had 5'T nucleotides preceding the TALEN binding site, 15 nucleotide long spacer regions, and a restriction enzyme cut site in the middle to perform RFLP assays for detecting potential mutations. Exon 2 (E2) TALEN left and right pair have 16 and 19 repeats while Exon 3 (E3) TALEN left and right pair have 17 and 18 repeats respectively. Figure 4.25 and Figure 4.26 shows the Exon 2 and Exon 3 regions, the TALEN pair binding sites and the primers that were used to amplify those regions.

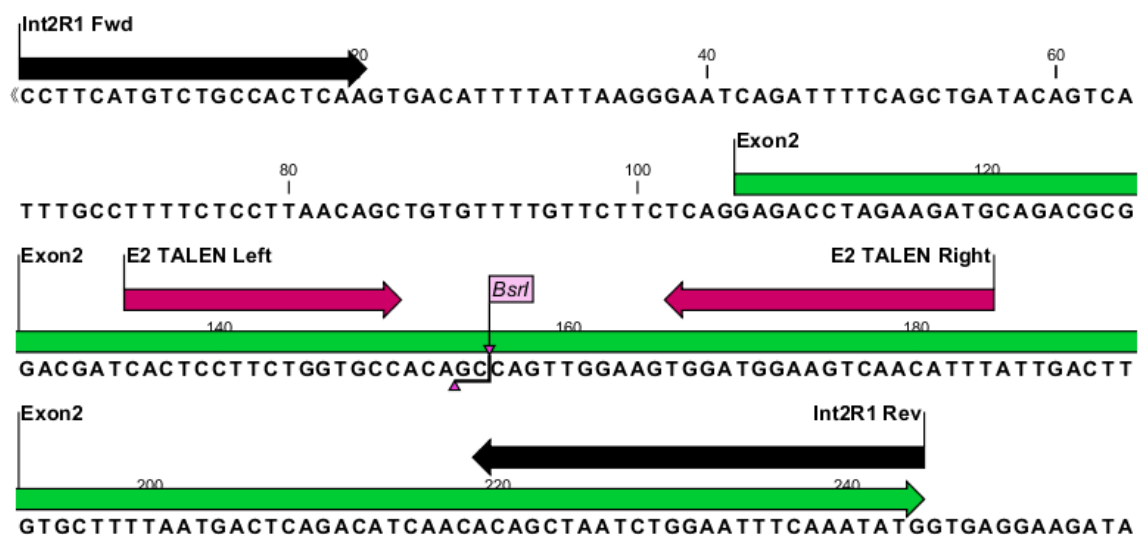


Figure 4.25 Binding sites of TALENs targeting Exon 2 of IL-7R gene. Magenta arrows indicate TALEN binding sites, black arrows show the primers that were used to amplify this region, the flag in the spacer region of TALEN pair shows the restriction enzyme BsrI cut site.

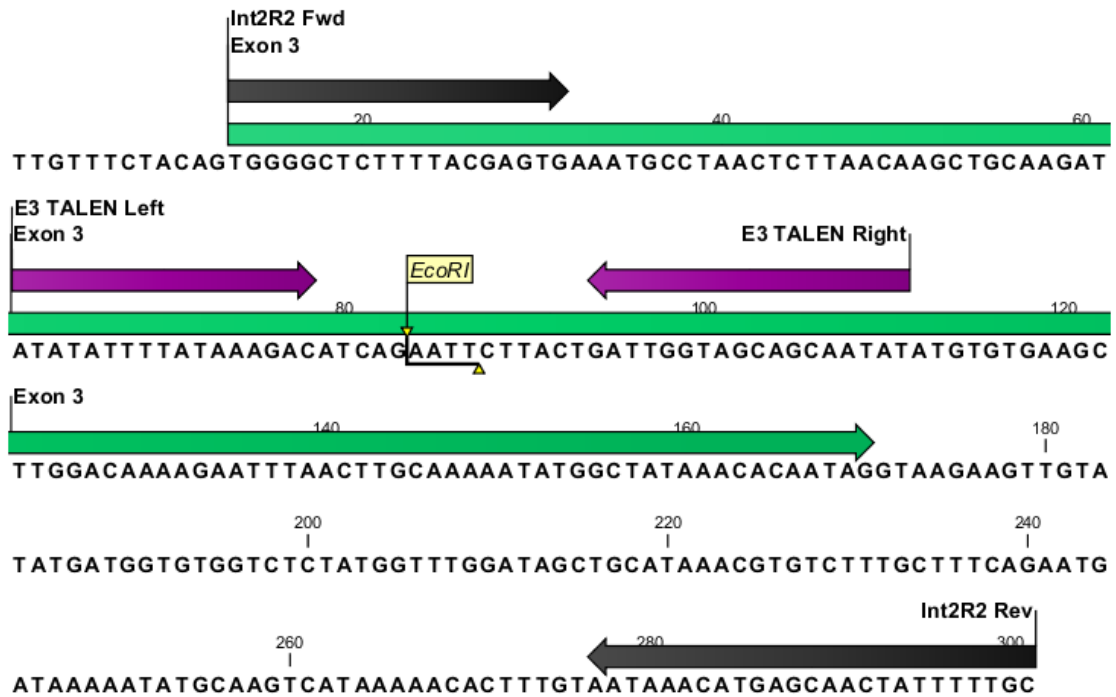


Figure 4.26 Binding sites of TALENs targeting Exon 3 of IL-7R gene. Purple arrows indicate TALEN binding sites, black arrows show the primers that were used to amplify this region, the flag in the spacer region of TALEN pair shows the restriction enzyme EcoRI cut site.

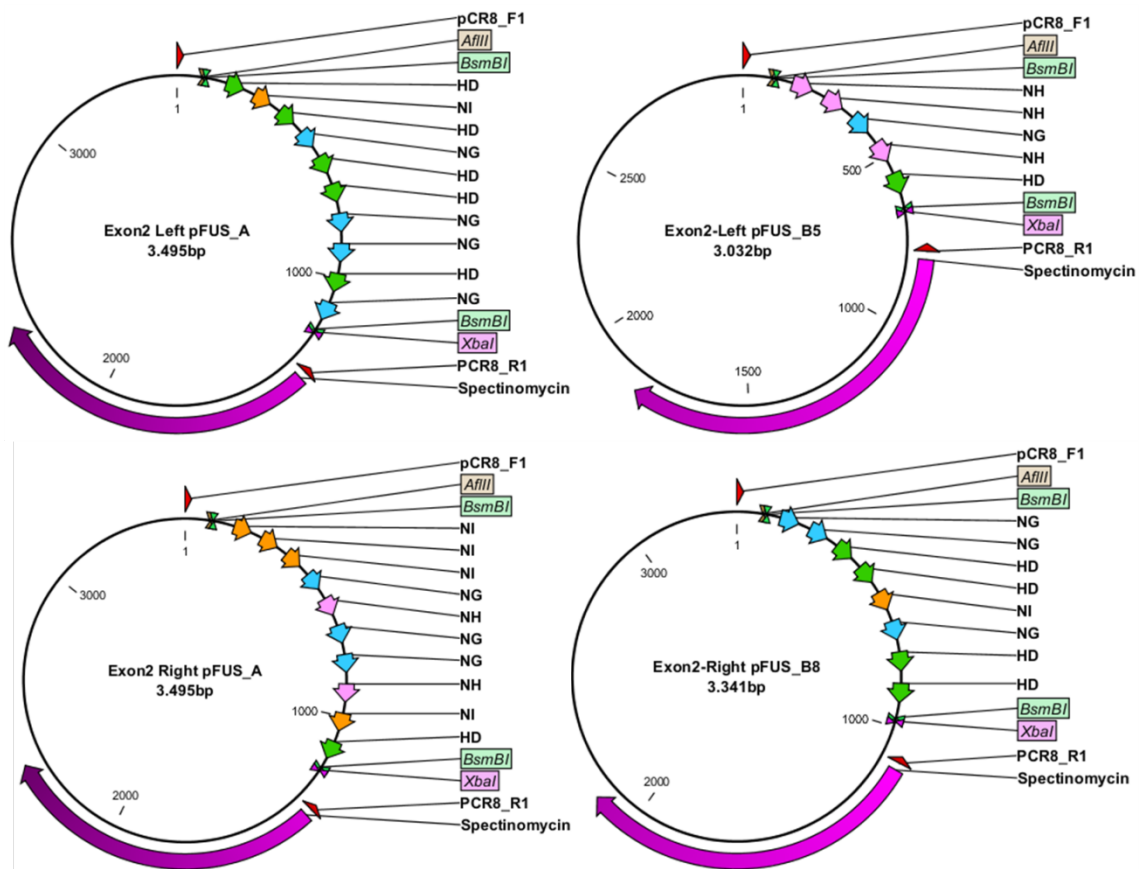


Figure 4.27 Plasmids maps of the Exon 2 targeting TALEN pair in pFUS_A and pFUS_B intermediary plasmids after Golden Gate reaction #1

We assembled the TALEN pairs using the Golden Gate Assembly kit. First 10 repeat modules were cloned into the pFUS_A plasmid and the remaining repeats except the last repeat were cloned into the pFUS_B plasmid of the respective number in Golden Gate reaction #1. After the transformation of the ligated array of repeats in the intermediary plasmids along with IPTG/X-gal blue-white screening, we chose the white colonies to perform colony PCR using the pCR8_F1 and PCR8_R1 primers. We expected to see bands sized 1200bp for pFUS_A that has 10 repeats and bands sized 700bp, 800bp, 900bp and 1000bp for pFUS_B5, pFUS_B6, pFUS_B7 and pFUS_B8 respectively. Along with the correct sized band a smear and a ladder effect caused by the repeats is also observed, consistent with the Golden Gate protocol. After choosing the correct colonies we isolated their DNA to do a confirmation digest with AflIII and XbaI. The 10 repeats in pFUS_A gives a 1048bp band while B5, B6, B7 and B8 give bands sized 523bp, 622bp, 720bp, and 820bp respectively. The plasmids maps for Exon 2 and Exon 3 TALEN pairs are shown in Figure 4.27 and Figure 4.28 and agarose gel images of colony PCR and control digest results for both pairs are shown in Figure 4.29.

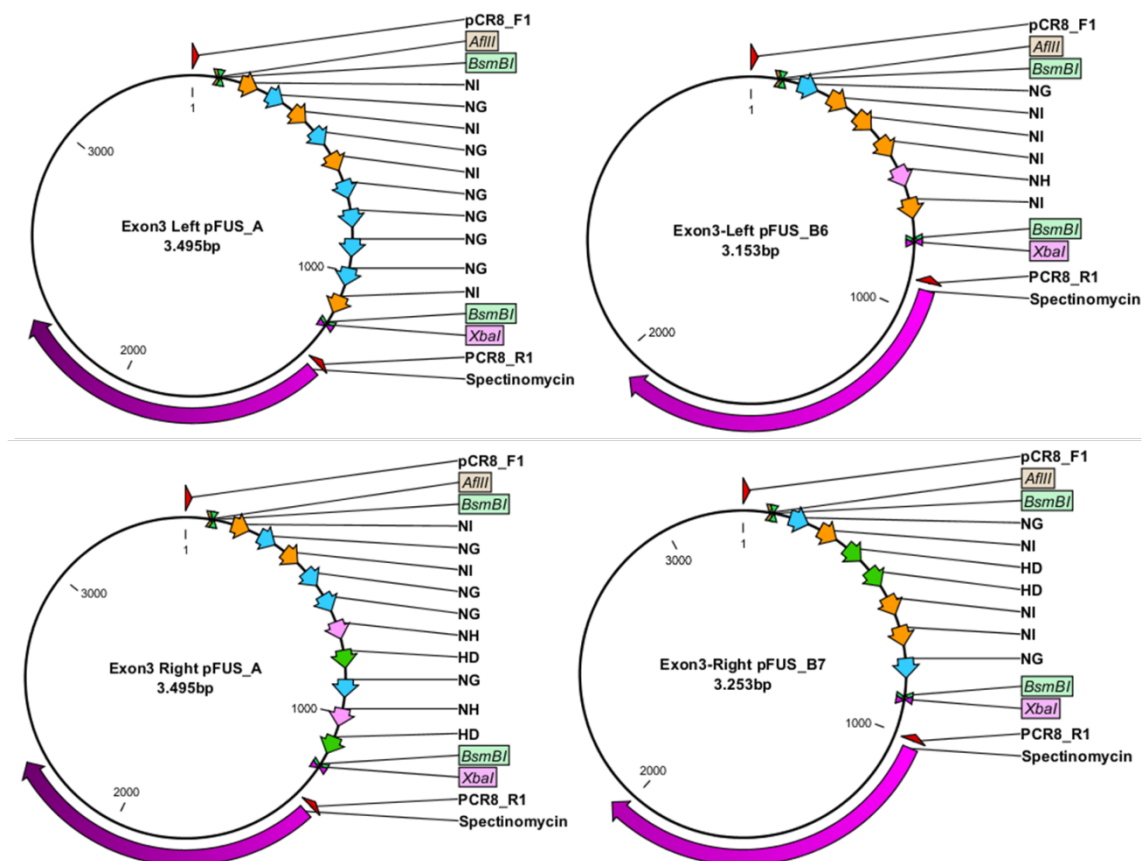


Figure 4.28 Plasmids maps of Exon 3 targeting TALEN pair in pFUS_A and pFUS_B intermediary plasmids after Golden Gate reaction #1

73

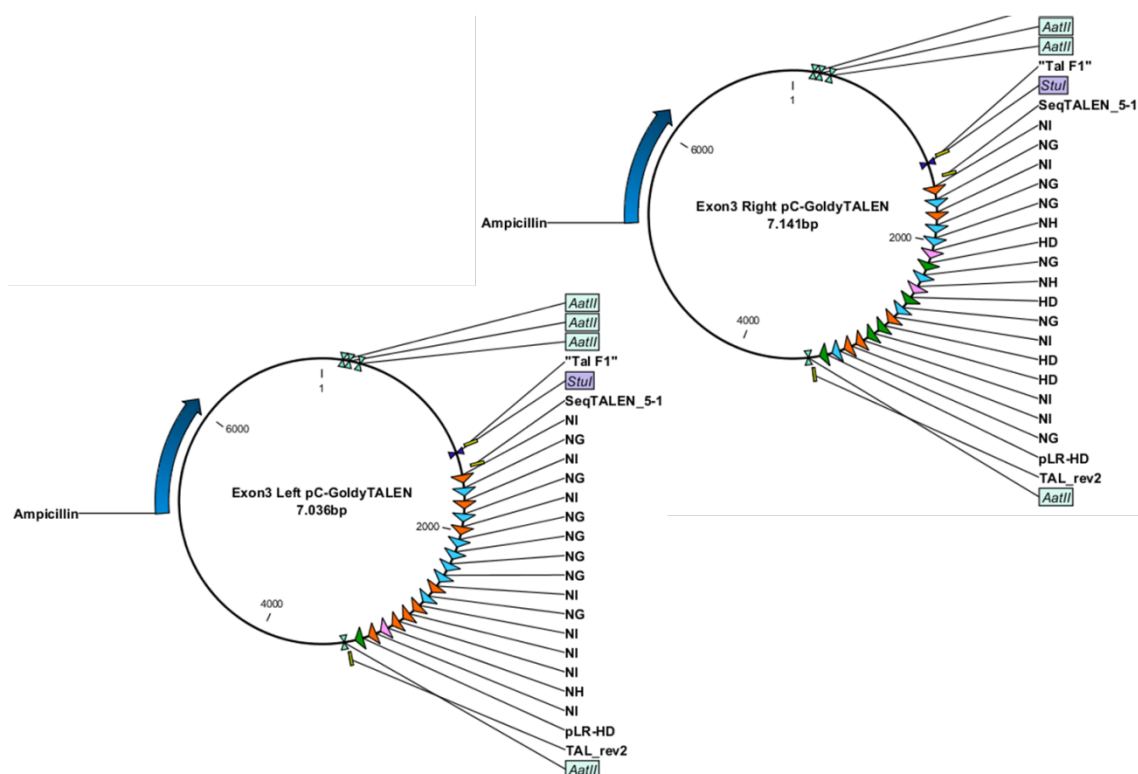


Figure 4.31 Plasmid maps for fully assembled Exon 3 targeting TALEN pair in their final expression vector pC-GoldyTALEN.

After the intermediary plasmids for TALEN repeats were assembled, we joined the pFUSA and pFUS_B plasmids together with the plasmid of the last repeat to the final expression plasmid named pC-GoldyTALEN in Golden Gate reaction #2. After transformation we searched for correct colonies performing colony PCR with the TAL_F1 and TAL_R2 primers, expecting to see bands around 1800bp-2000bp along with a smear and a ladder effect. After choosing 2 of the colonies that seemed to be correct, we isolated their DNA to perform a double digest with StuI and AatII. E2 Left TALEN should give a band sized 1826bp; E2 right 2132bp, E3 left 1928bp and lastly E3 right should be 2030bp. Figure 4.30 and Figure 4.31 show the plasmid maps for TALEN pairs in their final expression plasmid and Figure 4.32 shows the agarose images for colony PCR and control digest results for the final constructs after Golden Gate reaction #2.

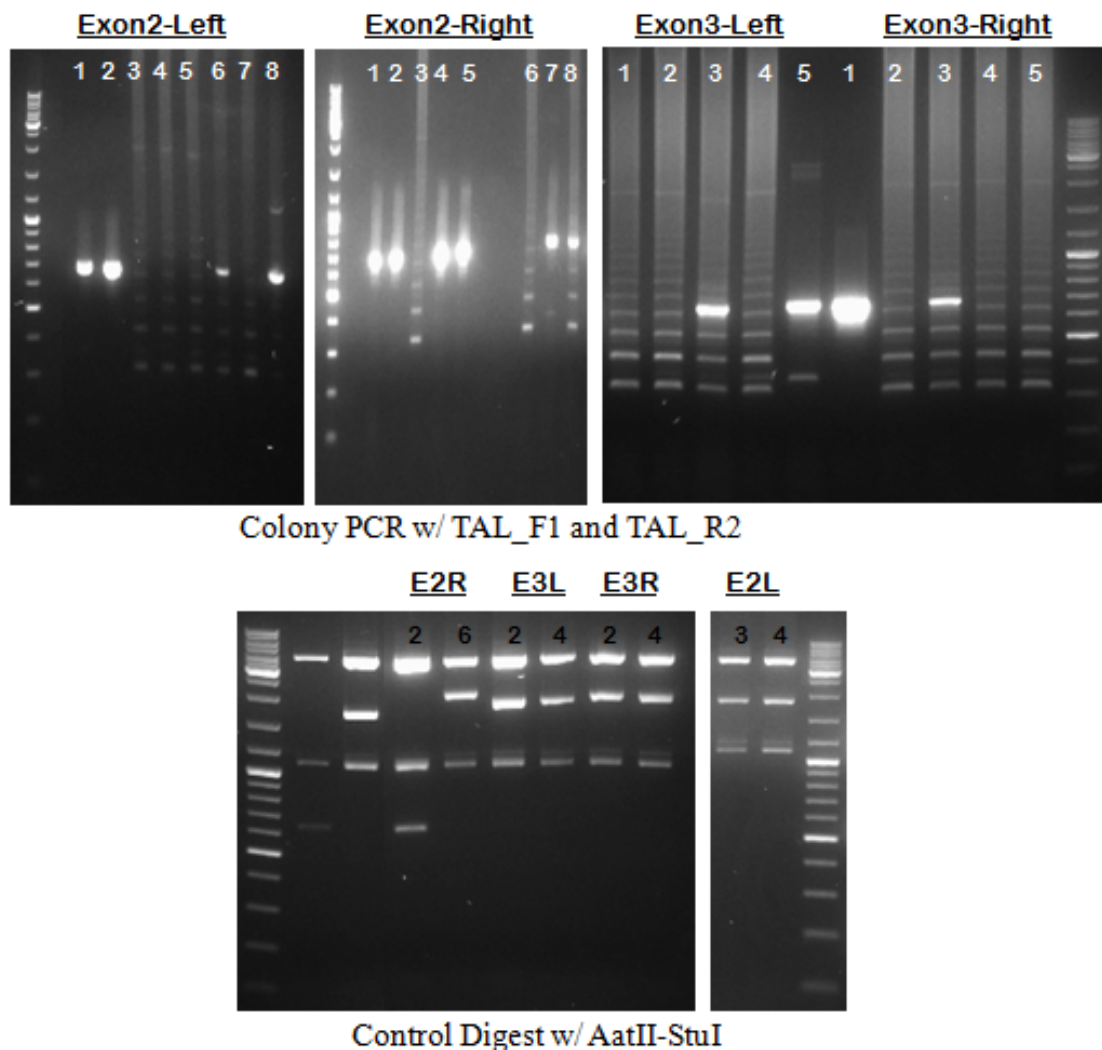


Figure 4.32 Agarose gel images of colony PCR and control digest results of both Exon 2 and Exon 3 targeting TALENs after Golden Gate reaction #2.

4.1.2.2. Expression of Exon2 and Exon3 TALEN pairs in RLM11 cells and detection of mutations

After the TALENs were assembled, we first transfected RLM11 cells with the Exon 2 and Exon 3 TALEN pairs separately to see if they function on their own, before starting with the contransfection experiments. After the transfection and TALEN expression were complete, we performed FACS analysis, expecting to see a decrease in IL-7R expression even in the pools, considering the fact that we directly targeted exons this time. Although we expected IL-7R levels of TALEN transfected cells to be significantly lower, it was possible that this effect would be partially due to electroporation and 32°C incubation. (Figure 4.33) To confirm these results we

performed the RFLP assay on the transfected cell pools and observed a clear uncut band for both; indicating that a quite large portion of the cells in the population were mutated (Figure 4.34).

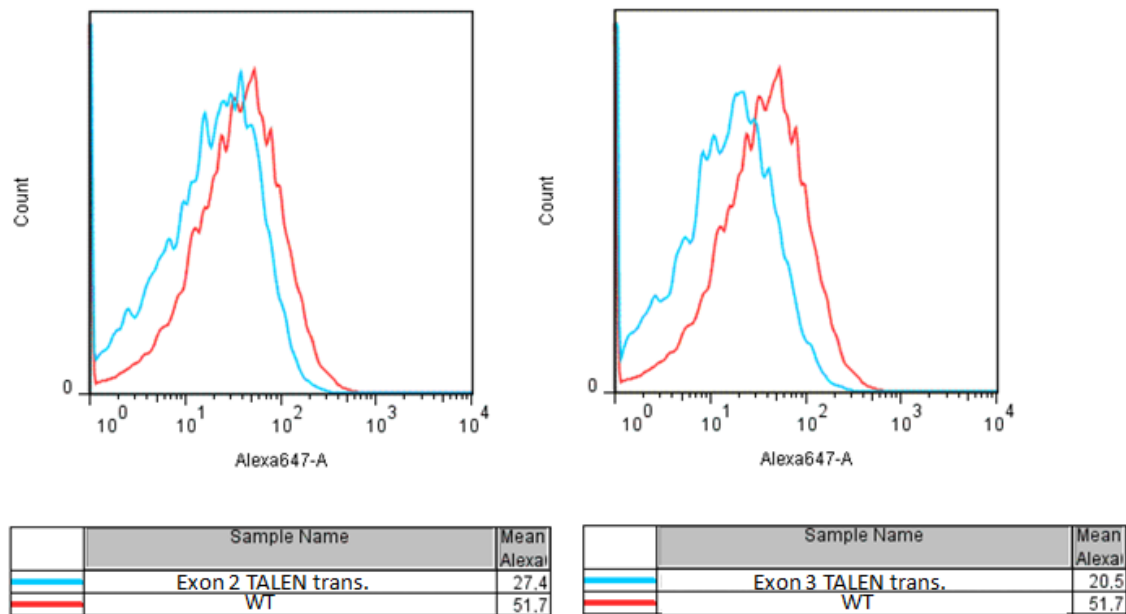


Figure 4.33 FACS analysis of Exon 2 and Exon 3 TALEN transfected RLM11 cells. Histogram for Alexa-647 represents surface IL-7R α expression. Red histogram represents WT while the blue ones represent transfected cells.

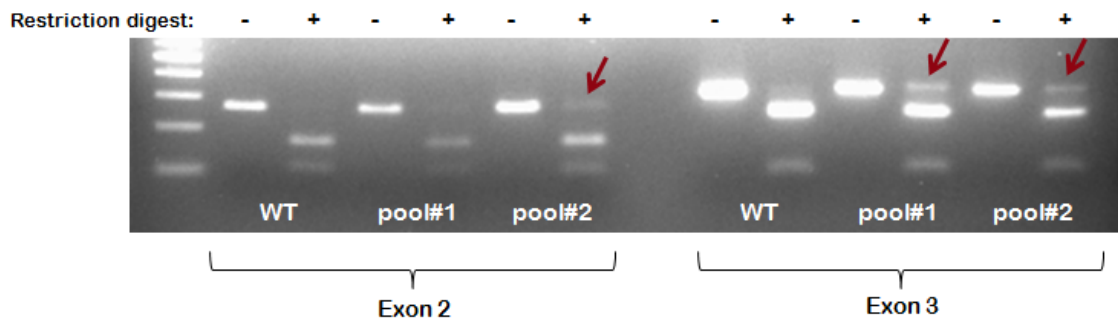


Figure 4.34 RFLP assay of the Exon 2 and Exon 3 TALEN transfected RLM11 cells. Exon 2 and Exon 3 were amplified using the primers shown in Figure 4.25 and Figure 4.26. Exon 2 PCR products were digested with BsrI enzyme while Exon 3 PCR products were digested with EcoRI. Pool#1 and Pool#2 stands for the pool of cells taken from the transfected samples at two different times; #1 right after 72h incubation including the dead cells after electroporation and #2 only the live cells that were grown after a few splits. Red arrows show the uncut bands caused by the presence of mutation.

4.1.2.3. A map of the donor plasmid targeting the ECR1 region between Exon 2 and 3 by homologous recombination

The donor plasmid to repair the IL-7R gene with homologous recombination after double strand breaks were induced by the E2 and E3 TALENs was previously generated by Gamze Günel in the Erman laboratory ⁷⁷. If homologous recombination is successful, this donor DNA will replace the Exon2-Intron2-Exon3 region of the IL7R gene and result in a new allele that has no Intron 2 and a fused Exon2 and Exon3. The protein encoded by this allele will be identical to the wild type IL7R protein. Exon 2 and Exon 3 PCR products were fused to each other using the strategy that was explained in Figure 3.1 of the Methods section and this amplicon was cloned into the pJET1.2/blunt plasmid using a CloneJET PCR cloning kit. After growing colonies from a glycerol stock and performing control digest with the XhoI-XbaI enzymes to make sure that the 600bp insert is still in this plasmid, we extracted the digested bands. Figure 4.35 shows the plasmid map of the construct and the control digests that were done to confirm the presence of Exon 2 and Exon 3 fusion PCR product.

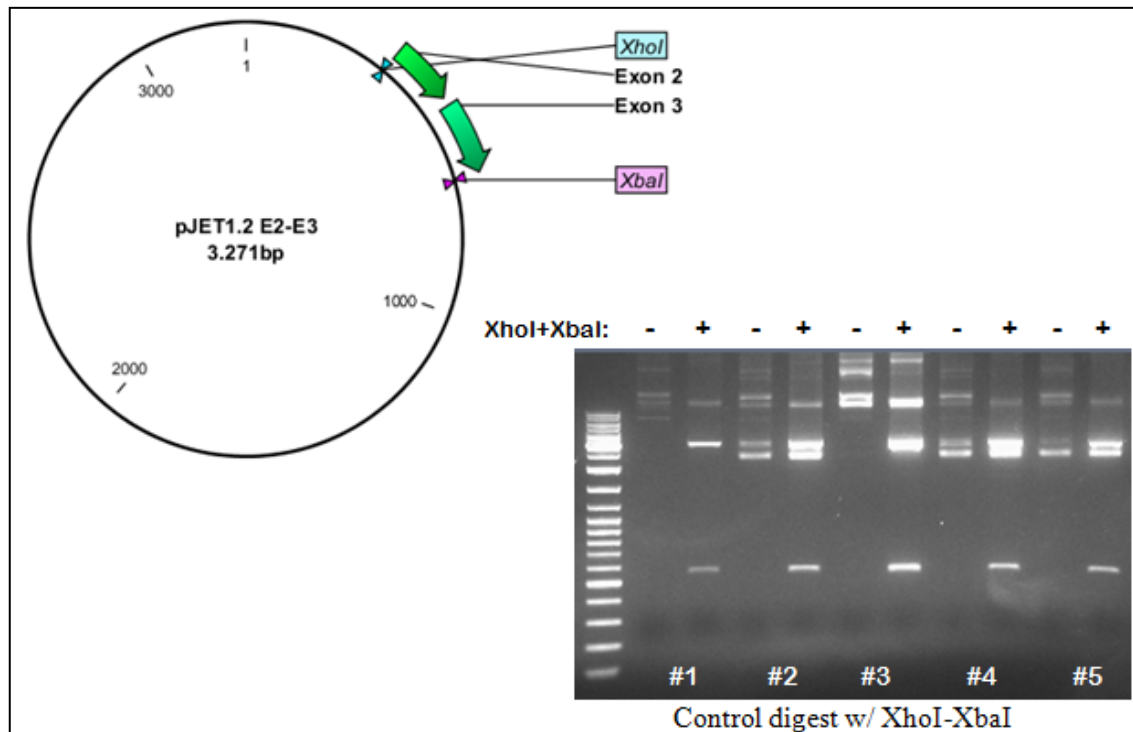


Figure 4.35 The plasmid map for the E2-E3 fusion product and the agarose gel image of the control digest.

4.1.2.4. Simultaneous transfection of the Exon2, Exon3 TALENs and the donor dsDNA in RLM11 cells

After confirming the activity of TALENs, we co-transfected RLM11 cells with both TALEN pairs and the homologous dsDNA donor. We used a 1:1 molar ratio for the total amount of TALENs and the donor dsDNA. Along with the main experiment we set up control experiments that we put the same amount of TALEN pairs with the co-transfection and complete the rest with mock DNA. After a 72h incubation at 32°C we isolated genomic DNA and performed the RFLP assay (Figure 4.36). In the RFLP assay, with cells transfected with only one of the pairs we detected uncut bands, indicating that these TALENs could mutate a significant proportion of the cells. Cells co-transfected with both TALEN constructs (4 plasmid in total) appeared to have slightly fainter uncut bands, which might indicate that both TALENs worked at the same time and removed the region in between. We also performed a PCR using forward primer of exon 2 and the reverse primer of exon 3 on the co-transfected sample and saw a band which corresponds to fused exon 2 and exon 3, however the sample only went through a dense medium split after the transfection and probably still contained the free donor DNA; we did not see this band in the pool after two more splits.

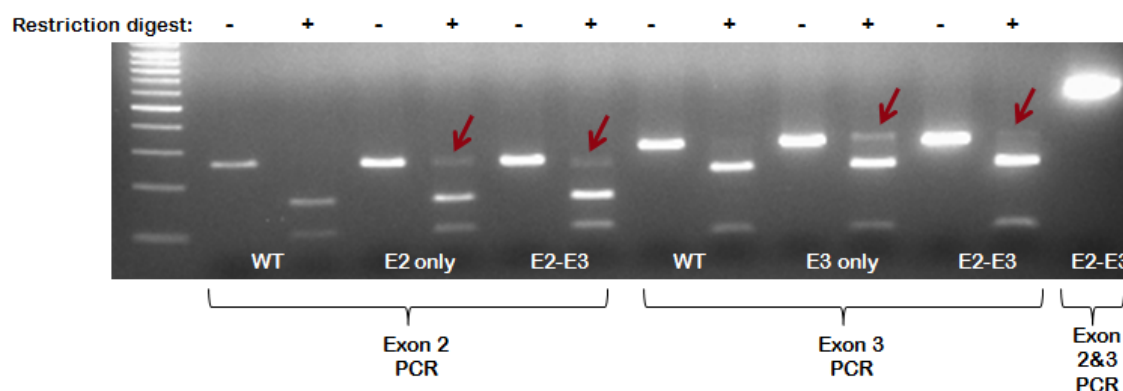


Figure 4.36 RFLP assay for Exon 2 and Exon 3 TALEN pair co-transfection. Exon 2 and Exon 3 were amplified using the primers shown in Figure 4.25 and Figure 4.26.

Exon 2 PCR products were digested with BsrI enzyme while Exon 3 PCR products were digested with EcoRI. The last lane was PCR amplified by the Exon 2 forward and Exon 3 reverse primers. Red arrows show the uncut bands caused by the presence of mutation.

After a few splits and making sure that the cells recovered their health after transfection, we performed FACS analysis on the samples to compare their IL-7R expression levels to WT expression. All of the transfected cells, the ones with Exon2 TALEN pair only, Exon3 TALEN pair only and the co-transfected ones had lower IL-7R expression compared to untransfected samples (Figure 4.37). All of these results were obtained from a pool of cells that still had wild type cells in majority. Experiments are continuing with the single cell colony screening.

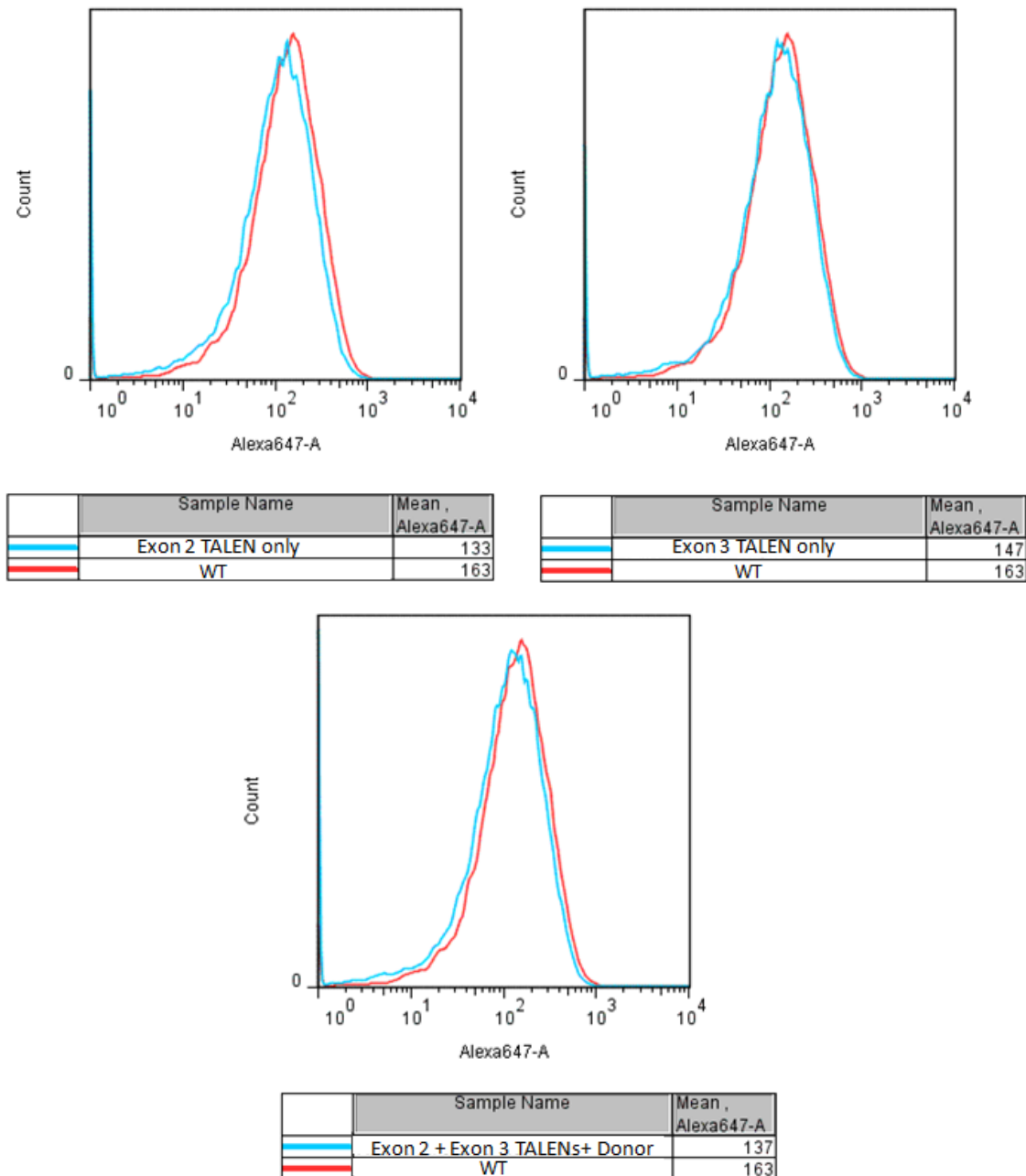


Figure 4.37 FACS analysis of Exon 2 and Exon 3 TALEN transfection and co-transfection in RLM11 cells. Histogram for Alexa-647 represents surface IL-7R α expression. Red histogram represents WT while the blue ones represent transfected cells.

4.2. Targeting GR gene

4.2.1. TALENs Targeting human GR gene to induce Knock-In and Knock-Outs

As we worked on the project that targets the transcription factor GR binding site of the IL-7R gene, we also approached the project from another side; we designed a TALEN pair against the glucocorticoid receptor gene itself to knock out the gene to see which genes would be affected from its absence; and by using the same TALEN pair and a homologous donor plasmid we designed, we aimed to insert Venus-YFP gene at the same site to make it fused to GR gene so that when they were expressed we could track GR protein's location within the cell when exposed to certain conditions.

To knock out the hGR gene, we designed TALENs that target a region near to translation start codon so that in case the cell fails to repair the gene after the TALEN pair induces a double strand break, non-homologous end joining would generate insertions or deletions that would cause frameshifts in the coding strand that would end up with a non-functional transcript. In case we provide a donor plasmid that is partly homologous to the region we induce mutation, through homologous recombination the changes we had done in the donor gene would transfer to the genome. For this experiment we designed a plasmid that has homologous sites to the genome on the left and right sides, and we aimed to insert puromycin resistance and Venus-YFP genes in the middle with no stop codons in between. Puromycin resistance would enable us to do selection among the ones that has the insert, and P2A sequence in between venus and puromycin genes would permit the translation of the transcript without detachment of the ribosome in between but have puromycin cleaved off from the venus-hGR complex right after the translation occurs so that the puromycin resistance in the cell would not interfere with their activity. The strategy to construct this donor plasmid is explained in Figure 3.2 of the Methods section and the strategy to knock-out or induce knock-ins in hGR is shown in Figure 4.38.

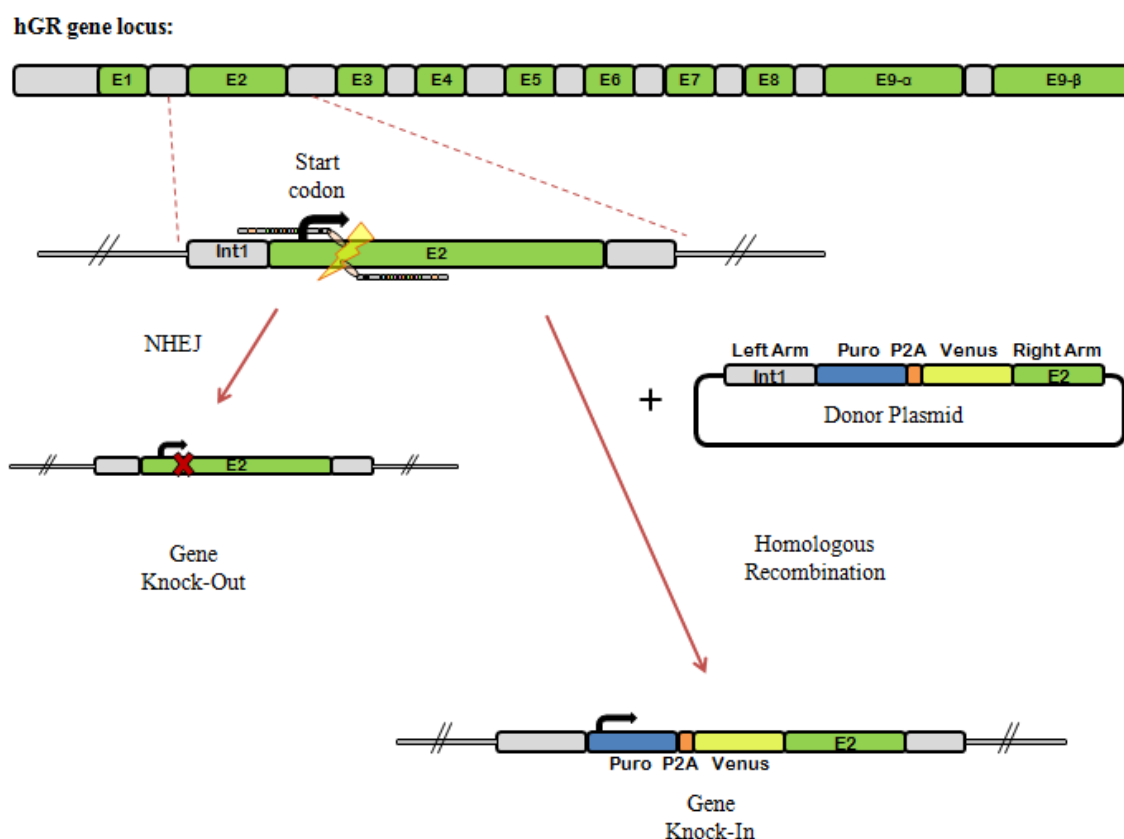


Figure 4.38 The strategy to knock out human glucocorticoid receptor using TALENs and insert Venus and puromycin resistance genes using homologous recombination.

4.2.2. Assembly of TALENs targeting translation start site of the hGR gene

We designed TALENs targeting hGR gene using the same methods and paying attention to similar things with the ones that was explained in section 4.1.1. The TALEN pair has 15 nucleotide spacer length in between, they are preceded by a 5'T nucleotide and there is a restriction enzyme recognition site in between to enable usage of RFLP assay to select mutants. The binding sites of the TALEN pair, the spacer region, and the start codon is shown on Figure 4.39.

Both left and right TALENs have 15 repeats, therefore in the Golden Gate reaction #1 we cloned first 10 repeats into pFUS_A array plasmid and four remaining ones in pFUS_B4. We had done IPTG/X-gal blue-white screening and had done colony PCR to the white colonies using pCR8_F1 and pCR8_R1 primers. The correct colonies gave 1200bp bands for 10 repeats and 600bp for 4 repeats. The plasmid maps and agarose gel images of the colony PCR products can be seen in Figure 4.40 and Figure 4.41.

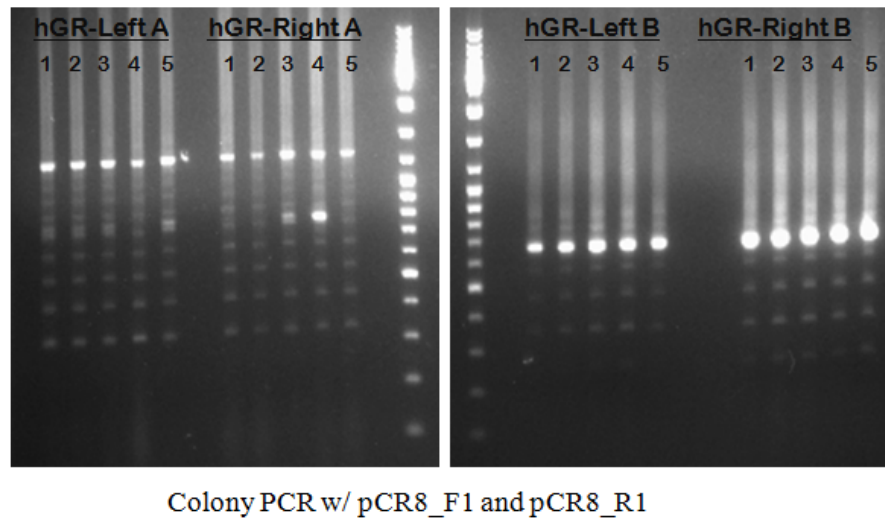


Figure 4.41 Agarose gel images of the colony PCR results of hGR TALENs after Golden Gate reaction #1

We joined the repeats in pFUS_A and pFUS_B plasmids along with the last repeats in pC-GoldyTALEN backbone with Golden Gate reaction #2. Using TAL_F1 and TAL_R2 primers we performed colony PCR expecting to see 1744bp bands along with a smear and ladder effect. Control digests were done with StuI and AatII, resulting in 1724bp band in between the ~1000bp and ~5000bp bands coming from pC-GoldyTALEN expression vector. The plasmid maps of fully assembled hGR TALEN pair is shown in Figure 4.42 and agarose gel images after colony PCR and their confirmation with AatII-StuI double digest is shown in Figure 4.43.

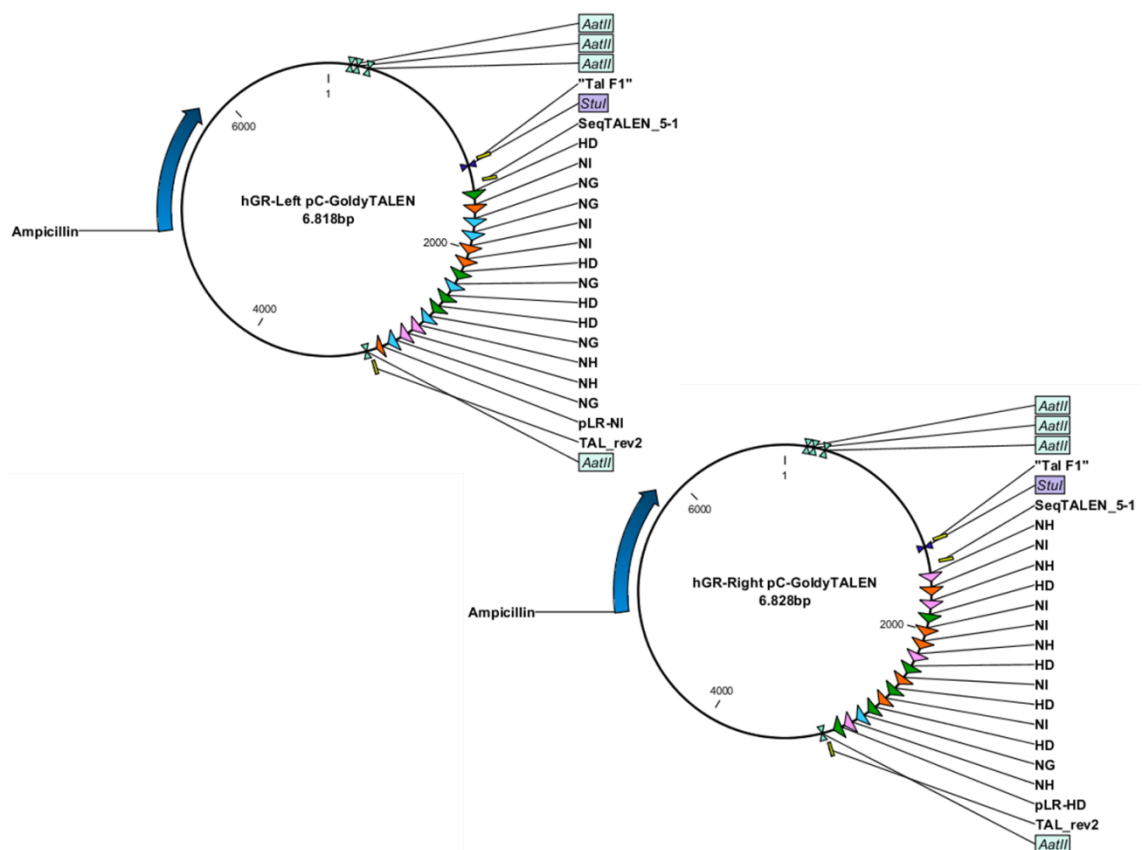


Figure 4.42: Plasmid maps of hGR left and right TALEN pair in pC-GoldyTALEN backbone.

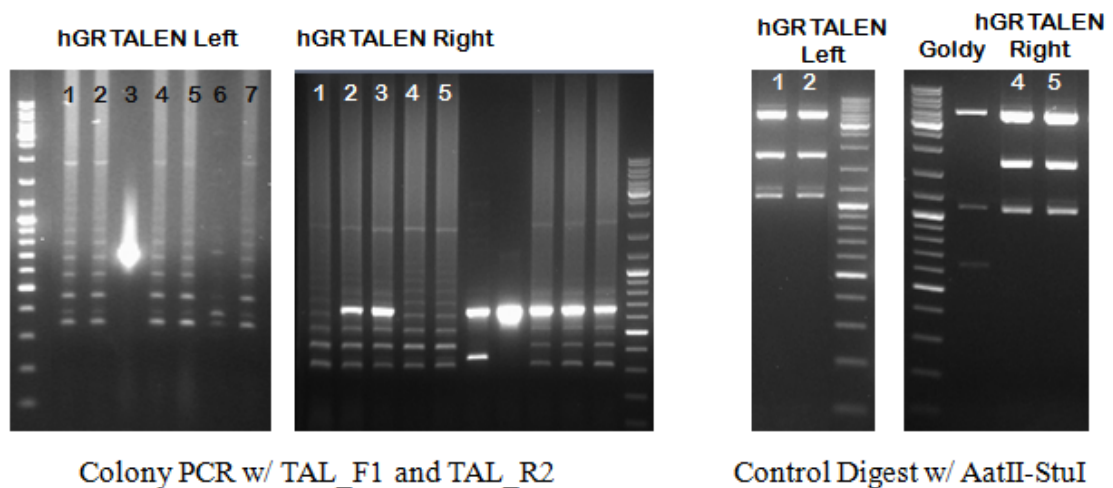


Figure 4.43: Colony PCR and double digest control for hGR TALEN pair after Golden Gate reaction #2.

4.2.3. Expression of hGR TALEN pair in HCT116 cells and detection of mutation

After the assembly of the TALEN pair was complete, we transfected adherent human colonic carcinoma cell line HCT116 with them using PEI transfection method. We incubated them at 32°C for 72 hours and extracted their genomic DNA to perform RFLP assay. As shown in Figure 4.44 there is a faint uncut band in transfected sample, indicating that the TALEN pair had induced mutation.

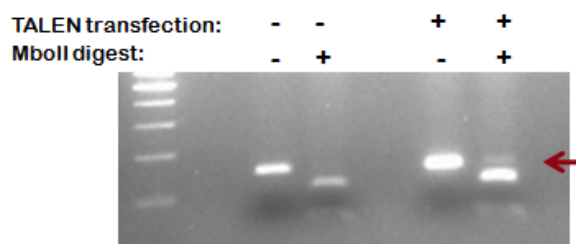


Figure 4.44 RFLP assay for hGR TALEN transfected HCT116 WT cells. Red arrow show the uncut band caused by the presence of mutation.

4.2.4. Construction of homologous donor plasmid to insert Venus gene into hGR endogenously

In order to induce gene knock-in at the translation start site of the glucocorticoid receptor gene we designed a donor plasmid that is homologous to the gene at both ends, and the middle part included the insert. Overall, there were four main different PCR products that were supposed to be joined; the left homologous arm, puromycin resistance gene, Venus-YFP gene and the right homologous arm. The strategy to assemble this product is explained in Figure 3.2 of the Methods section. Venus gene and the puromycin resistance gene have around 600bp long sequences therefore while designing left and right homologous arms we arranged them to have at least similar lengths to promote more accurate homologous recombination. Left arm was designed to have the whole promoter region before the start codon of the hGR gene so that when the two other genes were integrated into the genome at that point the start codon of the puromycin resistance gene would be the new start point of the translation. In order to have the expressed region of the hGR fused to the Venus gene the right homologous arm began with the start codon of the hGR, being planned to be attached right at the end of the Venus gene. The positions of the left and the right homologous arms on the hGR

gene are shown in Figure 4.45 and Figure 4.46. The forward primer of the left arm had *NheI* restriction enzyme cut site and the reverse primer of the right arm had *BamHI* cut site added on their 5' ends to ligate them to a plasmid when the whole product was assembled. We PCR amplified both regions with Phusion Polymerase enzyme for high fidelity and to avoid overhang ends. The agarose gel images of the PCR results are shown in Figure 4.47.

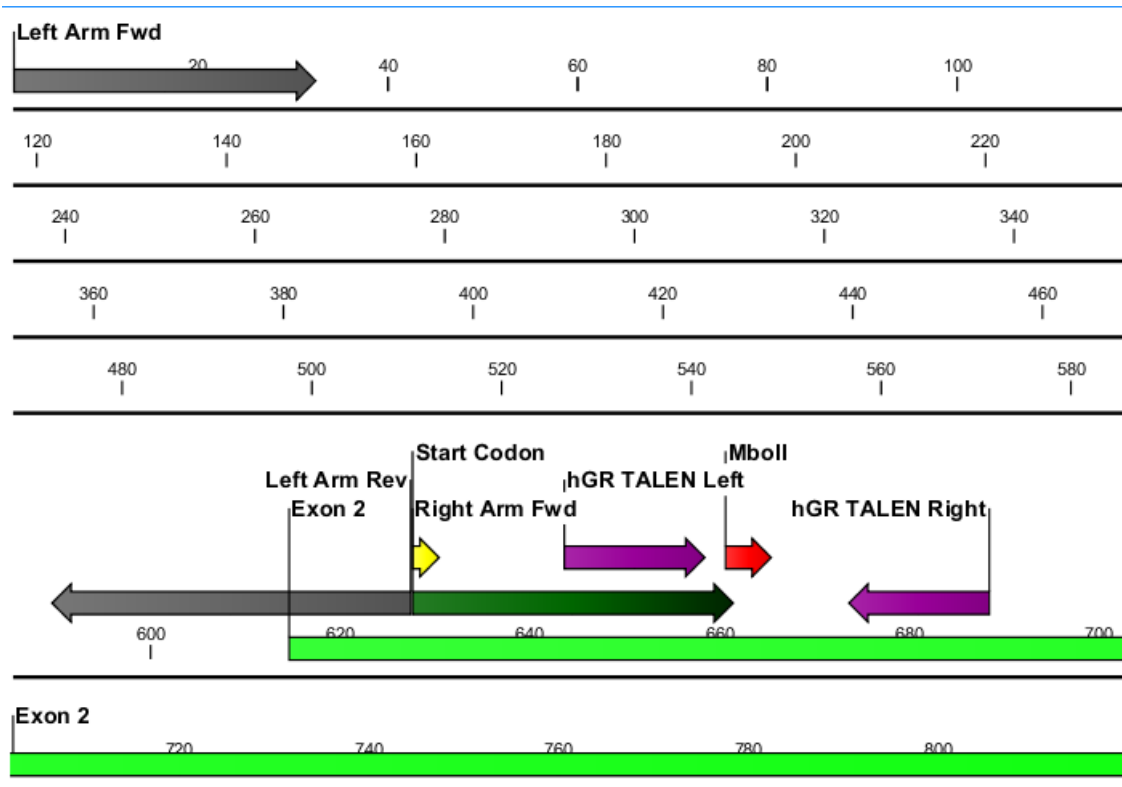


Figure 4.45 Left homologous arm PCR site on hGR gene locus. Gray arrows indicate the primers that are used to amplify the left homologous arm which is 627bp long. Light green arrow show the 2nd exon and the yellow arrow indicate the start codon. TALEN pair which are shown by purple arrows are at the downstream of the start codon.

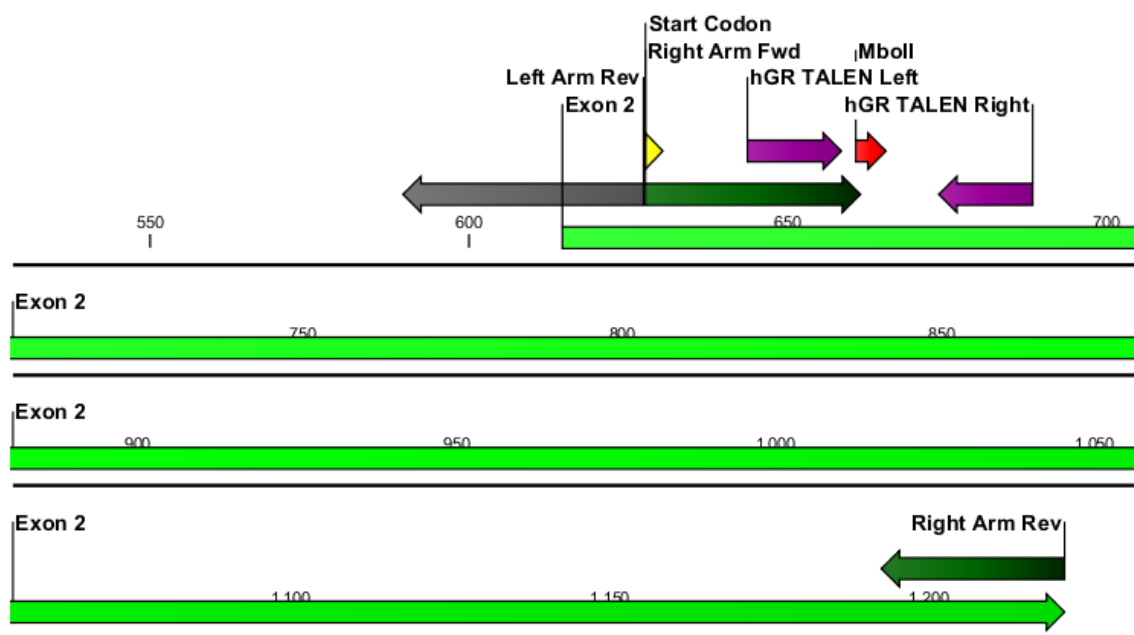


Figure 4.46 Right homologous arm PCR site on hGR gene locus. Dark green arrows indicate the primers that are used to amplify the right homologous arm which is 594bp long. Light green arrow show the 2nd exon and the yellow arrow indicate the start codon. TALEN pair which are shown by purple arrows are at the downstream of the start codon.

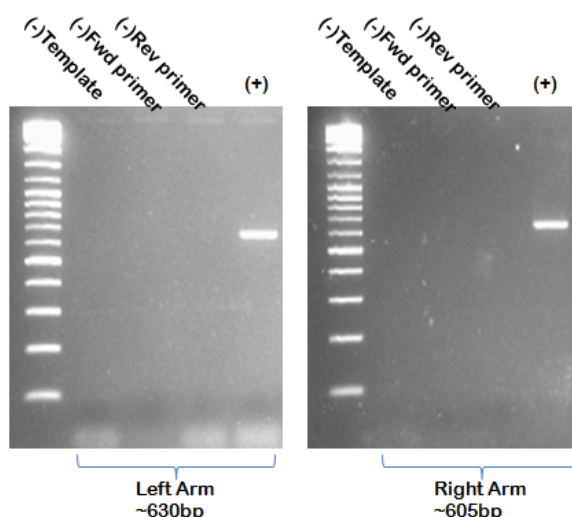


Figure 4.47 The optimized PCR results of left and right homologous arms of hGR gene. The regions were amplified from HCT116 genomic DNA extract.

We amplified the puromycin resistance gene from hAAVS-SA2A which was a donor plasmid used in another study⁴⁰. We added a ~20bp homologous site to the left arm's 3' end on the 5' end of the puromycin forward primer, and P2A sequence on the

5' end of the puromycin reverse primer. P2A is a self-cleaving peptide that has a ~60bp sequence which enables preceding and following genes' translation occur without ribosome detaching from the mRNA but as soon as the translation is complete the product is cleaved from this site, separating the two proteins from each other⁸⁰. We did not include the stop codon to the reverse primer of the puromycin resistance gene so that the translation would continue with the Venus gene without interruption but we put P2A sequence in between so that they would not be joined with Venus fluorescent protein. We amplified the Venus gene from mVenusC1, a plasmid modified from a YFP expressing vector⁸¹ and has no stop codon so that it could be attached to a protein that follows it. The venus forward primer had P2A sequence on its 5' end as the puromycin's reverse primer. The reverse primer of the venus gene was homologous to the right homologous arm's forward primer on its 5' end. The plasmid maps and the optimized PCR results for puromycin resistance gene and the venus gene are shown in Figure 4.48 and Figure 4.49.

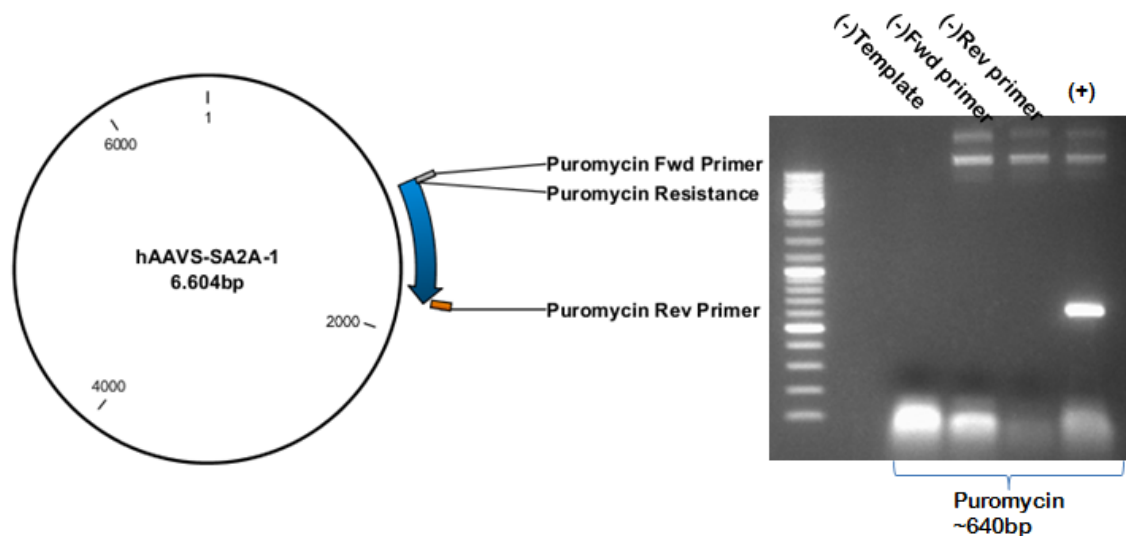


Figure 4.48 The plasmid map and optimized PCR result for puromycin resistance gene⁴⁰. Forward primer is homologous to the left homologous arm's reverse primer on the 5' end and the reverse primer includes P2A sequence on its 5' end.

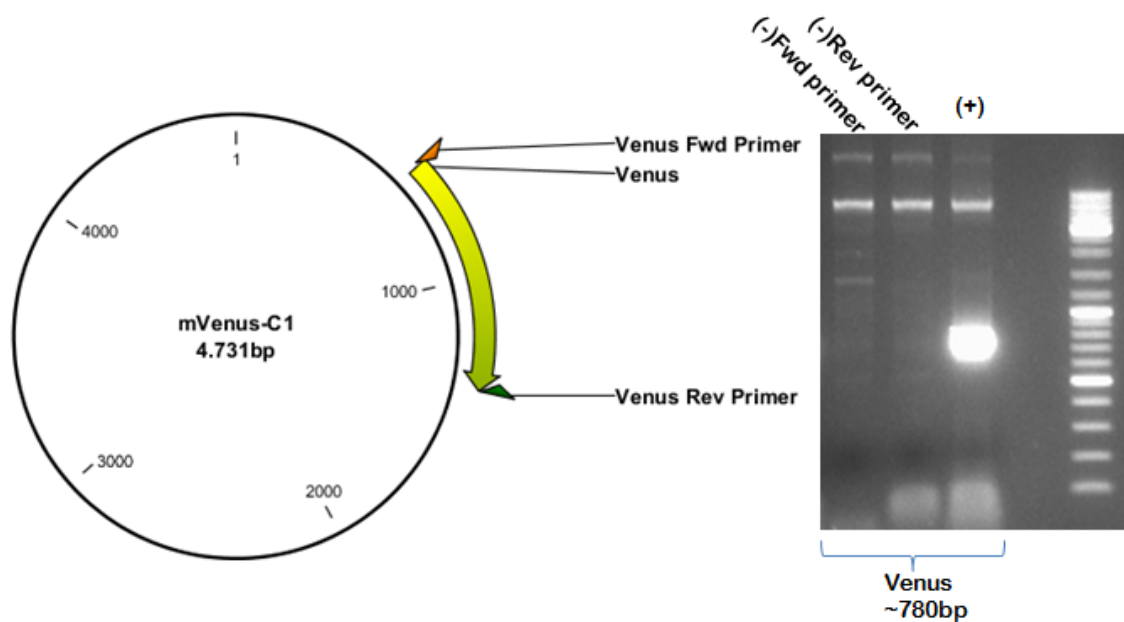


Figure 4.49 The plasmid map and optimized PCR result for Venus gene⁸¹. The forward primer includes P2A sequence on the 5' end and the reverse primer is homologous to the forward primer of the right arm on the 5' end.

After the PCR amplification for all four regions were complete we gel extracted them for purifying them from their primers and using the respective products as template we joined left arm and the puromycin resistance gene using left arm's forward primer and puromycin's reverse primer only; and we joined venus and right arm by using venus' forward primer and right arms reverse primer in a second PCR reaction. (Figure 4.50a) After gel extracting the correct bands we performed restriction digestion to the fused PCR products with Kpn2I enzyme; expecting to see bands around 750bp and 550bp for left arm-puromycin fusion, and around 750 and 650bp for venus-right arm fusion. (Figure 4.50b) Since we could not join the two products in a final PCR reaction instead we used Gibson Assembly protocol which was explained in Figure 3.3 to assemble the final product. (Figure 4.50c) Even though the ligation of the four PCR products are complete we could not yet clone it to a final plasmid, however the final design we aimed to construct is shown in Figure 4.51.

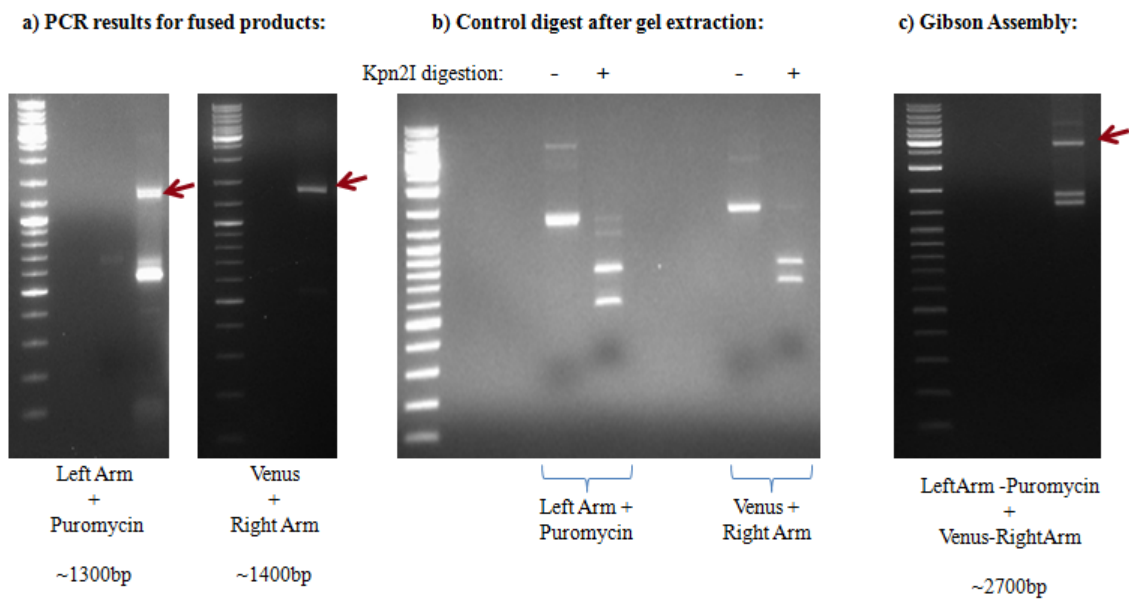


Figure 4.50 The experiments done to fuse PCR products of the hGR donor plasmid. The correct bands are shown with red arrows.

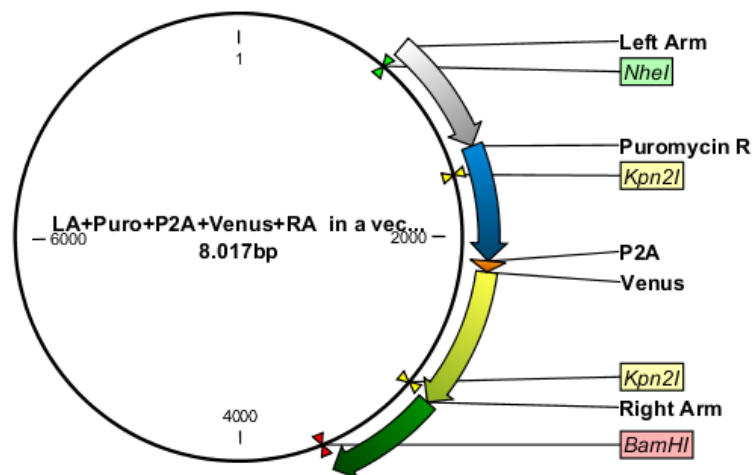


Figure 4.51 The plasmid map for the hGR donor construct. The gray arrow shows left homologous arm, the blue arrow shows puromycin resistance gene, the orange site that follows it is the P2A sequence, the yellow arrow indicates the venus gene and the green arrow is the right homologous arm, which is the second exon of the hGR gene.

5. DISCUSSION

Transcription activator-like effector (TALE) proteins from the plant pathogen *Xanthomonas* are recently discovered DNA binding proteins that are composed of highly conserved repeat modules in their DNA binding central domain. Each repeat module consists of 34 amino acids and they are polymorphic on the 12th and 13th amino acid residues which are called repeat variable di-residues (RVDs). The specificity of the DNA binding domain (DBD) is determined by a simple one RVD - one base code and modularity of this protein structure allows assembly of domains that can target any site in the genome. TALE nucleases (TALENs) were developed by the fusion of the catalytic domain of the FokI restriction enzyme to DBD of the TALE protein and now they are being used as tools for efficient genome modification ^{1,2}.

Interleukin-7 signaling is crucial for development, differentiation and survival of lymphocytes. The expression of IL-7 receptor in lymphocytes changes during development. Upstream promoter and enhancer regions of the IL-7R α gene locus contain binding sites for transcription factors that have roles in regulating this expression ^{63,65,66,70}. Also, the gene locus contains an evolutionarily conserved region with a potential silencer at the intronic region in between the 2nd and the 3rd exons. We mutated the binding sites of two transcription factors, glucocorticoid receptor (GR) and Notch, and we aimed to delete the entire 2nd intron to understand the roles of these sites in the regulation of IL-7R α gene expression in T lymphocytes. Glucocorticoids are immunosuppressive and anti-inflammatory agents that signal through glucocorticoid receptor and have various effects on growth, differentiation and function of lymphocytes ⁷². For this reason, along with the GR binding site in the IL-7R gene locus we targeted the translation start site of the glucocorticoid receptor gene itself, both for generating knockout cells that we could conduct experiments in the deficiency of GR

and to generate fluorescent GR proteins that could be visually tracked within the cell by inserting a Venus-YFP gene at the same locus and promoting their co-expression.

For our experiments we designed and constructed 5 pairs of TALENs and in all of them we paid attention to similar points, aiming to assemble the pairs that would bind the target with high specificity and would function efficiently. To avoid TALENs' off-target effects, they should only be expressed in the cells transiently; the presence of TALENs are no longer necessary after the double strand break (DSB) is induced and their continuous activity would not only increase the risk of other sites in the genome being targeted, but also would be deleterious to the mutations induced in the targeted site, preventing the generation of a stable mutant genotype. For this reason we cloned the TALEN repeats to eukaryotic expression plasmids and expressed these TALENs ectopically. Previous studies showed that the amino acid sequences flanking the DNA binding domain could affect the binding efficiency of the protein and the distance between the DBD and FokI cleavage domain in the C-terminal could be an important factor in their dimerization and catalytic activity. Goldy TALEN backbone was generated to optimize TALEN activity with truncations in the N and C terminal regions and was shown to be very efficient^{34,42}. Therefore, we used Goldy TALENs as our expression plasmids in all of the constructs. As RVDs, we used NI, NG, HD for adenine, thymine, and cytosine base recognition respectively. In Notch TALEN experiments NN was used for binding the G nucleotide due to its high binding efficiency compared to other guanine binding RVDs. However, to avoid possible off-target effects we switched to NH RVD for TALEN constructions that were generated later on since NH was reported to be more specific than NN even though it is less efficient in binding^{9,15}. The length of the spacer region in between two TALEN binding sites is another factor that affects the dimerization and catalytic activity of the FokI domain, and recent studies show that a 15-16 bp long spacer is the optimum spacer length for most efficient TALEN activity⁸². While constructing the TALENs in these experiments we took these parameters into account.

Detection of the TALEN induced mutations was challenging. After we obtained a TALEN introduced pool of cells, we performed restriction fragment length polymorphism (RFLP) assays to determine the efficiency of TALEN activity. The principle of this strategy is to have a restriction enzyme cut site in the targeted region.

Induced deletions and insertions (INDELs) disrupt the region, destroy the restriction enzyme (RE) recognition sequence and block REs from digesting this target from amplified PCR products, appearing as an uncut band. However, not all of the targeted sites conveniently had restriction enzyme cut sites right in the middle, and considering that most of the mutations induced would span a small site, it is very likely that some of the mutations could not be detected because they did not affect the RE recognition site. For instance, we could only detect a very faint uncut band in hGR TALEN transfected HCT116 cells (Figure 4.44). The recognition site for the MboII enzyme only spanned a region on the left side of the spacer; therefore, we could not detect INDELs appearing on the right side (Figure 4.39). These kinds of parameters make it difficult to determine the efficiency of TALEN activity and detect the presence of mutations.

We used RLM11 cells, a CD4 single positive thymoma cell line which has high IL-7R expression levels on their membranes for observing the effects of induced mutations, assuming that it could easily be detected by the changes on the expression phenotype. However, the IL-7R expression level comparison of TALEN induced cell pools did not yield reliable results because it took a long time for the cells to recover after electroporation and 32°C incubation for three days. We subjected cells to transient hypothermia (32°C incubation) after TALEN transfection because it was shown that this could increase the mutation efficiency both by zinc finger nucleases and TALENs^{10,83} probably by increasing the folding efficiency of the proteins and slowing down the progression of the cell cycle. Another problem was the difficulty of sorting the small populations of cells due to the low efficiency of genomic mutations. When we compared IL-7R expression levels by flow cytometry right after the transfection and 32°C incubation we saw that GFP control plasmid transfected cells also had much lower IL-7R levels in most of the experiments (Figure 4.9). When we waited for a few splits for cells to recover and eliminate the effects of transfection, we observed that the transfected pools did not have any significant difference in their IL-7R expression levels compared to wild type cells. It was likely that we lost various types of mutations during the splits, and in case the mutant cells had any disadvantage of growing, the WT cells within the pool could take over the population. For this reason, we decided to grow single cell colonies from both GR and Notch binding site targeting TALEN introduced pools so that we could detect their specific mutations and directly correlate a mutation with a phenotype.

We screened for GR transcription factor binding site mutated RLM11 single cell colonies by the RFLP assay and we found 5 colonies that give an uncut band upon digestion. We analyzed their mutations by DNA sequencing and confirmed that all the mutations had occurred in the GR binding site (Figure 4.13). When we compared their IL-7R expression levels to WT cells we realized that most of the cells had higher expression levels than the WT in repeated experiments except for the 10th colony (Figure 4.14). These results were unexpected because glucocorticoid treatment is used to induce IL-7R expression and GR was shown to bind to this location to positively regulate the IL-7R expression in cells that do not express IL-7R⁷³. It is possible that GR signaling does not have a role in the regulation of IL-7R gene expression in late stage thymocytes that express IL-7R stably and maybe even has suppressing effects even though we cannot confirm this only with the present results. The 10th colony on the other hand probably has lower IL-7R levels because it had a large deletion (19bp) in the targeted site, which is likely to disrupt a neighboring TF binding site of a positive regulator. We repeated this experiment with 3B4 cells which do not normally express IL-7R, but can be stimulated by glucocorticoid treatment; however, we had to discontinue the experiments due to the low efficiency of transfections. We think that disruptions induced on the GR binding site we targeted would prevent the complete activation of IL-7R expression in these cells.

It was more challenging to screen for Notch transcription binding site mutated single cell colonies than the GR mutated ones. If we take uncut bands in their RFLP analysis results as an indicator of mutation rate, around 20-25% of the GR binding sites in the pool genome were mutated while this percentage was only around 5% for the Notch TALENs. For that reason while screening for mutated single cells we eliminated a portion of them by flow cytometric analysis selecting only colonies with lower IL-7R expression levels compared to WT cells (Figure 4.18). We assumed that the mutated colonies would be among these IL-7R^{low} cells. We performed an RFLP assay on the selected colonies and found only two of these to be mutated (Figure 4.19). While doing these experiments we realized that growing colonies from a single cell reduces their health and recovery takes a long time. Most of the cells we selected at first probably had lower expression only due to the growth conditions. Upon further culturing, most clones recovered their IL-7R expression. However, a few clones still had significantly lower IL7R expression levels than the WT pool. We analyzed the mutation status of these

clones by DNA sequencing. We predicted that these clones might have a mutation that could not be detected by RFLP. However, only the 35th colony had mutations in the Notch binding site and all the others, including the ones with lowest IL-7R expressions (clone 30th and 74th) had intact GR binding sites (Figure 4.21). The 74th colony was the most interesting one among the colonies because it had a knock-out phenotype; however, both RFLP and the sequencing results indicated that it had no mutations. We hypothesized that this could result from a very large deletion in the locus that prevents the amplification of the region by PCR. Therefore, we also PCR amplified the region with primers that span a larger area in the locus. However, we still observed only the WT bands (Figure 4.23). We do not yet know the reason for the loss of expression in these cells. The phenotype we observe might be due to an even larger deletion that we could not detect or it could be because of an off-target activity of the TALEN pair, resulting in the disruption of an alternative locus important for IL7R expression. With these experiments, we show that the Notch site mutated colony #35 and #49 have significantly reduced IL7R expression levels compared to WT cells, yet we do not know why IL-7R expression in the other few cells had decreased or whether there were any mutants among the ones that had normal IL-7R levels.

Another problem about the interpretation of the results we obtained with both the GR and Notch binding site targeting experiments was the unknown ploidy of the RLM11 cancer cells. Even though we grew single cells to eliminate the WT phenotype, most of the mutations induced seemed to be monoallelic and even clones with two types of mutations also contained an intact WT. The GR binding site mutated cell colony #15 had two types of deletions (Figure 4.13) and even though it was barely visible in RFLP results there we also observed a WT sequence, which was confirmed by sequencing analysis. In a similar way, colony #35 from Notch binding site targeted cells had one insertion and one deletion (Figure 4.21). Yet, a third WT sequence appeared in both the RFLP and sequencing analyses. This could be due to a low frequency WT population in the clone; a WT cell that possibly arising from adhered to the mutant cell when single cells were separated into different wells. This WT contaminant might have started to divide much later and still might be appearing as a small portion of the clone. Alternatively, these results could be due to presence of at least three alleles, which will be confirmed by SKY karyotype analysis. If the latter is the case, then the mutant colonies with one type of mutation have at least two of their alleles intact, and due to

this reason their phenotype might not be getting affected by the mutation. Nonetheless, at least two of the alleles in colonies #11 and #15 of the GR site experiments were mutated. The same situation is valid for the 35th colony of Notch experiments.

For the experiments where we targeted the 2nd intron of the IL-7R gene, we showed that even with co-expression of two TALEN pairs we could induce double strand breaks. However, in that cell pool it is likely that these mutations were not induced in the same cells or it is also possible that even when they were in the same cells, one pair might have induced the DSB after the other was repaired and sealed back. Nonetheless, we observed by the RFLP assay that in comparison to control experiments performed with only one of the TALEN pairs, the uncut band was fainter for the co-transfection experiments (Figure 4.36). This might either indicate that a small portion of the mutated cells contained DSBs induced simultaneously and that the locus in between the two target sites was deleted, or that the DSBs were repaired by homologous recombination due to the presence of donor DNA. We performed PCR with the forward primer of the exon 2 and the reverse primer of the exon 3 in order to detect if the truncated product. Even though we observed the band one split after the incubation of the cells, it most likely belonged to the donor DNA; a few splits later we did not observe any bands in the pool. We have also grown single cell colonies from the first pool and currently screening for the correct mutations.

For the last part of the study we aimed to insert a Venus-YFP gene to the translation start site of the glucocorticoid receptor in order to promote their co-expression endogenously and create a model that we could track GR protein activities within the cell visually. Even though the homologous recombination rate is smaller than NHEJ, when TALENs are used with selectable markers such as antibiotic resistance, the detection of mutation becomes much simpler. There are examples of studies using reporter systems in which a selectable marker is expressed only when mutations are generated at the TALEN target site ^{40,84}. For that reason, while designing the donor plasmid we also included a puromycin resistance gene in between the left homology arm and the Venus gene so that when the donor DNA is integrated into the genome through homology directed repair, the first translation start codon would belong to the puromycin resistance gene. Because there were no stop codons in between them, Venus and GR would be translated along with the puromycin gene. We inserted a small P2A

sequence in between the Venus and puromycin genes which would permit the translation of the mRNA without detachment of the ribosome in the middle, but would have puromycin polypeptide cleaved off from the Venus-hGR fusion protein immediately after the translation. So far, we confirmed that the TALEN pair targeting the site is functional (Figure 4.44). However, even though we constructed the final donor product, it is yet to be cloned into a plasmid for transfection experiments.

6. CONCLUSION

In this study, we aimed to induce various types of mutations in the IL-7R α and GR gene loci using TALEN technology. Firstly, we targeted the IL-7R gene enhancer and intronic sites in order to understand the functional role of two transcription factor sites in the enhancer region and a possible silencer site in the intronic region. The literature indicates that the enhancer of the IL-7R is bound by several transcription factors. We targeted two of these sites, GR and Notch. We constructed TALEN proteins that target the binding sites of these transcription factors in the IL-7R enhancer and examined the effects on gene expression. We generated single cell clones that contain at least one mutated IL-7R allele in the murine RLM11 cell line. We found that targeting the Notch site downregulated IL-7R expression. On the other hand, targeting the GR binding site did not significantly downregulate IL-7R expression in this cell line. In addition to these transcription factor binding sites, we targeted an evolutionarily conserved region in intron 2 of the IL-7R gene. For this purpose, we constructed TALEN pairs against exon 2 and exon 3 of the gene to induce simultaneous double strand breaks at these sites. We demonstrated that these TALENs can actively generate mutations in these sites. However, we have not yet completed the generation of a mutant IL-7R allele that has deleted the ECR in intron 2.

In the second part of the study, we targeted the GR gene itself. In order to knock out the activity of this gene, we constructed TALEN pairs targeting the translation initiation site. These TALENs can effectively generate INDEL mutations in this locus, resulting in frameshift mutations. We also targeted the same locus by homology directed repair to insert a Venus fluorescent protein gene in frame with the GR cDNA. Targeting of this gene locus is not yet complete. However, a Venus-YFP-GR fusion protein encoded by the endogenous GR gene locus will generate a model that will allow us to track GR activity visually. In conclusion we generated 10 different constructs

encoding five different TALEN pairs targeting either the coding region or enhancer sequences of two immunologically important genes. These TALENs are functional and can efficiently introduce INDEL mutations in the genomes of tissue culture cell lines. In this thesis we demonstrate that TALEN induced genome editing is an effective tool to study the functional significance of cis regulatory regions in the genomes of tissue culture cell lines.

REFERENCES

1. Boch, J. & Bonas, U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu. Rev. Phytopathol.* **48**, 419–36 (2010).
2. Chen, K. & Gao, C. TALENs: Customizable Molecular DNA Scissors for Genome Engineering of Plants. *J. Genet. Genomics* **40**, 271–279 (2013).
3. Moscou, M. J. & Bogdanove, A. J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009).
4. Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009).
5. Mak, A. N.-S. N.-S., Bradley, P., Cernadas, R. A., Bogdanove, A. J. & Stoddard, B. L. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* **335**, 716–9 (2012).
6. Deng, D. *et al.* Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* **335**, 720–3 (2012).
7. Jankele, R. & Svoboda, P. TAL effectors: tools for DNA Targeting. *Brief. Funct. Genomics* (2014). doi:10.1093/bfgp/elu013
8. Weber, E., Gruetzner, R., Werner, S., Engler, C. & Marillonnet, S. Assembly of designer tal effectors by golden gate cloning. *PLoS One* **6**, (2011).
9. Cong, L., Zhou, R., Kuo, Y., Cunniff, M. & Zhang, F. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat. Commun.* **3**, 968 (2012).
10. Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–8 (2011).
11. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39**, e82 (2011).

12. Schmid-Burgk, J. L., Schmidt, T., Kaiser, V., Höning, K. & Hornung, V. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nat. Biotechnol.* **31**, 76–81 (2013).
13. Mercer, A. C., Gaj, T., Fuller, R. P. & Barbas, C. F. Chimeric TALE recombinases with programmable DNA sequence specificity. *Nucleic Acids Res.* **40**, 11163–11172 (2012).
14. Christian, M. L. *et al.* Targeting G with TAL Effectors: A Comparison of Activities of TALENs Constructed with NN and NK Repeat Variable Di-Residues. *PLoS One* **7**, (2012).
15. Streubel, J., Blücher, C., Landgraf, A. & Boch, J. TAL effector RVD specificities and efficiencies. *Nat. Biotechnol.* **30**, 593–5 (2012).
16. Gaj, T., Gersbach, C. a & Barbas, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397–405 (2013).
17. Urnov, F. D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646–651 (2005).
18. Soldner, F. *et al.* Generation of isogenic pluripotent stem cells differing exclusively at two early onset parkinson point mutations. *Cell* **146**, 318–331 (2011).
19. Perez, E. E. *et al.* Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **26**, 808–816 (2008).
20. Liu, Q., Segal, D. J., Ghiara, J. B. & Barbas, C. F. Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 5525–5530 (1997).
21. Moehle, E. a *et al.* Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 3055–60 (2007).
22. Cathomen, T. & Joung, J. K. Zinc-finger nucleases: the next generation emerges. *Mol. Ther.* **16**, 1200–1207 (2008).
23. Sun, N. & Zhao, H. Transcription activator-like effector nucleases (TALENs): A highly efficient and versatile tool for genome editing. *Biotechnol. Bioeng.* (2013). doi:10.1002/bit.24890
24. Li, T. *et al.* TAL nucleases (TALNs): Hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* **39**, 359–372 (2011).
25. Wood, A. J. *et al.* Targeted genome editing across species using ZFNs and TALENs. *Science* **333**, 307 (2011).

26. Li, T., Liu, B., Chen, C. Y. & Yang, B. TALEN utilization in rice genome modifications. *Methods* (2014). doi:10.1016/j.ymeth.2014.03.019
27. Christian, M., Qi, Y., Zhang, Y. & Voytas, D. F. Targeted mutagenesis of *Arabidopsis thaliana* using engineered TAL effector nucleases. *G3 (Bethesda)*. **3**, 1697–705 (2013).
28. Liu, J. *et al.* Efficient and Specific Modifications of the *Drosophila* Genome by Means of an Easy TALEN Strategy. *J. Genet. Genomics* **39**, 209–215 (2012).
29. Huang, P. *et al.* Heritable gene targeting in zebrafish using customized TALENs. *Nat. Biotechnol.* **29**, 699–700 (2011).
30. Lei, Y. *et al.* Efficient targeted gene disruption in *Xenopus* embryos using engineered transcription activator-like effector nucleases (TALENs). *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17484–9 (2012).
31. Sung, Y. H. *et al.* Knockout mice created by TALEN-mediated gene targeting. *Nat. Biotechnol.* **31**, 23–4 (2013).
32. Wefers, B. *et al.* Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 3782–7 (2013).
33. Tesson, L. *et al.* Knockout rats generated by embryo microinjection of TALENs. *Nat. Biotechnol.* **29**, 695–696 (2011).
34. Carlson, D. F. *et al.* Efficient TALEN-mediated gene knockout in livestock. *Proc. Natl. Acad. Sci.* **109**, 17382–17387 (2012).
35. Chen, S. *et al.* A large-scale in vivo analysis reveals that TALENs are significantly more mutagenic than ZFNs generated using context-dependent assembly. *Nucleic Acids Res.* **41**, 2769–2778 (2013).
36. Reyon, D. *et al.* FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol.* **30**, 460–465 (2012).
37. Mussolino, C. *et al.* TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. *Nucleic Acids Res.* **42**, 6762–6773 (2014).
38. Joung, J. K. & Sander, J. D. TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* **14**, 49–55 (2013).
39. Maresca, M., Lin, V. G., Guo, N. & Yang, Y. Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. *Genome Res.* **23**, 539–46 (2013).
40. Hockemeyer, D. *et al.* Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* **29**, 731–4 (2011).

41. Zu, Y. *et al.* TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat. Methods* **10**, 329–31 (2013).
42. Bedell, V. M. *et al.* In vivo genome editing using a high-efficiency TALEN system. *Nature* (2012). doi:10.1038/nature11537
43. Ma, S. *et al.* Highly Efficient and Specific Genome Editing in Silkworm Using Custom TALENs. *PLoS One* **7**, (2012).
44. Wefers, B., Ortiz, O., Wurst, W. & Kühn, R. Generation of targeted mouse mutants by embryo microinjection of TALENs. *Methods* (2014). doi:10.1016/j.ymeth.2014.01.002
45. Namen, A. E. *et al.* Stimulation of B-cell progenitors by cloned murine interleukin-7. *Nature* **333**, 571–573 (1988).
46. Von Freeden-Jeffry, U. *et al.* Lymphopenia in interleukin (IL)-7 gene-deleted mice identifies IL-7 as a nonredundant cytokine. *J. Exp. Med.* **181**, 1519–1526 (1995).
47. Khaled, A. R. & Durum, S. K. Lymphocide: cytokines and the control of lymphoid homeostasis. *Nat. Rev. Immunol.* **2**, 817–830 (2002).
48. Al-Rawi, M. A. A., Mansel, R. E. & Jiang, W. G. Interleukin-7 (IL-7) and IL-7 receptor (IL-7R) signalling complex in human solid tumours. *Histol. Histopathol.* **18**, 911–923 (2003).
49. Jiang, Q. *et al.* Cell biology of IL-7, a key lymphotrophin. *Cytokine Growth Factor Rev.* **16**, 513–533 (2005).
50. Cosenza, L., Gorgun, G., Urbano, A. & Foss, F. Interleukin-7 receptor expression and activation in nonhaematopoietic neoplastic cell lines. *Cell. Signal.* **14**, 317–325 (2002).
51. González-García S, García-Peydró M, A. J. and T. M. in *Notch Regul. Immune Syst.* 47–73 (Springer Berlin Heidelberg, 2012).
52. Ceredig, R. & Rolink, A. G. The key role of IL-7 in lymphopoiesis. *Semin. Immunol.* **24**, 159–64 (2012).
53. Peschon, J. J. *et al.* Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice. *J. Exp. Med.* **180**, 1955–1960 (1994).
54. Corcoran, A. E., Riddell, A., Krooshoop, D. & Venkitaraman, A. R. Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. *Nature* **391**, 904–907 (1998).
55. Johnson, K. *et al.* Regulation of immunoglobulin light-chain recombination by the transcription factor IRF-4 and the attenuation of interleukin-7 signaling. *Immunity* **28**, 335–45 (2008).

56. Kikuchi, K., Lai, A. Y., Hsu, C.-L. & Kondo, M. IL-7 receptor signaling is necessary for stage transition in adult B cell development through up-regulation of EBF. *J. Exp. Med.* **201**, 1197–1203 (2005).
57. Mackall, C. L., Fry, T. J. & Gress, R. E. Harnessing the biology of IL-7 for therapeutic application. *Nat. Rev. Immunol.* **11**, 330–342 (2011).
58. Hong, C., Luckey, M. A. & Park, J.-H. H. Intrathymic IL-7: the where, when, and why of IL-7 signaling during T cell development. *Semin. Immunol.* **24**, 151–8 (2012).
59. Yu, Q., Erman, B., Park, J.-H., Feigenbaum, L. & Singer, A. IL-7 receptor signals inhibit expression of transcription factors TCF-1, LEF-1, and RORgammat: impact on thymocyte development. *J. Exp. Med.* **200**, 797–803 (2004).
60. Singer, A., Adoro, S. & Park, J.-H. Lineage fate and intense debate: myths, models and mechanisms of CD4- versus CD8-lineage choice. *Nat. Rev. Immunol.* **8**, 788–801 (2008).
61. Bradley, L. M., Haynes, L. & Swain, S. L. IL-7: Maintaining T-cell memory and achieving homeostasis. *Trends Immunol.* **26**, 172–176 (2005).
62. DeKoter, R. P., Lee, H. J. & Singh, H. PU.1 regulates expression of the interleukin-7 receptor in lymphoid progenitors. *Immunity* **16**, 297–309 (2002).
63. DeKoter, R. P. *et al.* Regulation of the interleukin-7 receptor alpha promoter by the Ets transcription factors PU.1 and GA-binding protein in developing B cells. *J. Biol. Chem.* **282**, 14194–14204 (2007).
64. Egawa, T., Tillman, R. E., Naoe, Y., Taniuchi, I. & Littman, D. R. The role of the Runx transcription factors in thymocyte differentiation and in homeostasis of naive T cells. *J. Exp. Med.* **204**, 1945–57 (2007).
65. Vosshenrich, C. a J. *et al.* A thymic pathway of mouse natural killer cell development characterized by expression of GATA-3 and CD127. *Nat. Immunol.* **7**, 1217–24 (2006).
66. Kerdiles, Y. M. *et al.* Foxo1 links homing and survival of naive T cells by regulating L-selectin, CCR7 and interleukin 7 receptor. *Nat. Immunol.* **10**, 176–184 (2009).
67. Miller, M. L. *et al.* Basal NF- κ B controls IL-7 responsiveness of quiescent naïve T cells. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7397–402 (2014).
68. Kopan, R. *et al.* Notch Signaling. (2014). doi:10.1101/cshperspect.a011213
69. Radtke, F. *et al.* Deficient T cell fate specification in mice with an induced inactivation of Notch1. *Immunity* **10**, 547–58 (1999).

70. González-García, S. *et al.* CSL-MAML-dependent Notch1 signaling controls T lineage-specific IL-7R $\{\alpha\}$ gene expression in early human thymopoiesis and leukemia. *J. Exp. Med.* **206**, 779–791 (2009).
71. Smoak, K. A. & Cidlowski, J. A. Mechanisms of glucocorticoid receptor signaling during inflammation. *Mech. Ageing Dev.* **125**, 697–706 (2004).
72. Ashwell, J. D., Lu, F. W. & Vacchio, M. S. Glucocorticoids in T cell development and function*. *Annu. Rev. Immunol.* **18**, 309–345 (2000).
73. Lee, H.-C., Shibata, H., Ogawa, S., Maki, K. & Ikuta, K. Transcriptional regulation of the mouse IL-7 receptor alpha promoter by glucocorticoid receptor. *J. Immunol.* **174**, 7800–7806 (2005).
74. Franchimont, D. *et al.* Positive effects of glucocorticoids on T cell function by up-regulation of IL-7 receptor alpha. *J. Immunol.* **168**, 2212–2218 (2002).
75. Ligon, D. L. *et al.* CD8 Lineage-specific Regulation of Interleukin-7 Receptor Expression by the Transcriptional Repressor Gfi1. *J. Biol. Chem.* **287**, 34386–34399 (2012).
76. Silverstone, A., Sun, L., Witte, O. N. & Baltimore, D. Biosynthesis of murine terminal deoxynucleotidyltransferase. *J. Biol. Chem.* **255**, 791–796 (1980).
77. Gunal, G. Molecular biological investigation of interleukin-7 receptor gene expression in T lymphocytes. (2009).
78. Temiz, S. T. Generation of Xanthomonas derived TALE proteins that inhibit gene transcription. (2013).
79. Longo, P. A., Kavran, J. M., Kim, M. S. & Leahy, D. J. Transient mammalian cell transfection with polyethylenimine (PEI). *Methods Enzymol.* **529**, 227–240 (2013).
80. Osborn, M. J. *et al.* A picornaviral 2A-like sequence-based tricistronic vector allowing for high-level therapeutic gene expression coupled to a dual-reporter system. *Mol. Ther.* **12**, 569–574 (2005).
81. Koushik, S. V, Chen, H., Thaler, C., Puhl, H. L. & Vogel, S. S. Cerulean, Venus, and VenusY67C FRET reference standards. *Biophys. J.* **91**, L99–L101 (2006).
82. Ma, A. C., Lee, H. B., Clark, K. J. & Ekker, S. C. High efficiency In Vivo genome engineering with a simplified 15-RVD GoldyTALEN design. *PLoS One* **8**, e65259 (2013).
83. Doyon, Y. *et al.* Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat. Methods* **8**, 74–79 (2011).
84. Kim, H. *et al.* Magnetic Separation and Antibiotics Selection Enable Enrichment of Cells with ZFN/TALEN-Induced Mutations. *PLoS One* **8**, (2013).

APPENDIX

APPENDIX A: Chemicals Used in the Study

<u>Chemicals</u>	<u>Supplier Company</u>
Acetic Acid	Sigma, Germany
Agarose	Sigma, Germany
Ampicillin sodium salt	Cellgro, USA
Boric acid	Molekula, UK
Calcium chloride	Sigma, Germany
Distilled water	Milipore, France
DMEM	Gibco, USA
DMSO	Sigma, Germany
EDTA	Sigma, Germany
Ethanol	Sigma, Germany
Ethidium bromide	Sigma, Germany
Fetal Bovine Serum (FBS)	Lonza, Switzerland
Glycerol	Sigma, Germany
HBSS	Gibco, USA
Hydrochloric Acid	Sigma, Germany
Isopropanol	Sigma, Germany
LB Agar	BD, USA
LB Broth	Sigma, Germany
L-glutamine	Hyclone, USA
Liquid nitrogen	Karbogaz, Turkey
2-mercaptoethanol	Sigma, Germany
Penicillin-Streptomycin	Sigma, Germany
PIPES	Sigma, Germany

PBS	Sigma, Germany
RNase A	Roche, Germany
RPMI 1640	Gibco, USA
SDS	Sigma, Germany
Sodium Azide	Amresco,USA
Spectinomycin	Sigma, Germany
Tris base	Sigma, Germany
Trypan Blue	Fluca, Germany
Trypsin-EDTA	Gibco, USA

APPENDIX B: Equipment Used In the Study

Equipment	Company
Autoclave	Priorclave, UK
Balance	Sartorius, BP221S, Germany
	Schimidzu, Libror EB-3200 HU, Japan
Centrifuge	Hitachi, Sorvall RC5C Plus, USA
	Eppendorf, 5415D, Germany
	Eppendorf, 5418R, Germany
	Beckman Coulter, Allegra®X-15R, USA
CO ₂ Incubator	Binder, Germany
Deepfreeze	-80°C, Forma, Thermo Electron Corp., USA
	-20°C, Bosch, Germany
Distilled Water	Millipore, Elix-S, France
Electrophoresis Apparatus	Biorad Inc., USA
	VWR, USA
Electroporator	Invitrogen, Neon Transfection Systems, USA
Flow Cytometer	BD FACS Canto, USA
Gel Documentation	Biorad, UV-Transilluminator 2000, USA
Heater	Thermomixer Comfort, Eppendorf, Germany
Hematocytometer	Hausser Scientific, Blue Bell Pa., USA
Ice Machine	Scotsman Inc., AF80, USA
Incubator	Memmert, Modell 300, Germany
Laminar Flow	Kendro Lab. Prod., Heraeus, HeraSafe HS12, Germany
Liquid Nitrogen Tank	Taylor-Wharton, 3000RS, USA
Magnetic Stirrer	Stuart™, SB162, UK

Microliter Pipettes	Gilson, Pipetman, France
Microscope	Olympus IX70, Japan Olympus CK40, Japan
Microwave Oven	Bosch, Germany
pH meter	Mettler Toledo, S220 SevenCompact™ pH/Ion, USA
Refrigerator	Bosch, Germany
Shaker Incubator	New Brunswick Sci., Innova 4330, USA
Spectrophotometer	Amersham Biosciences, UK
Thermocycler	Eppendorf, Mastercycler Gradient, Germany
Vortex	Velp Scientifica, Italy

APPENDIX C: DNA Molecular Weight Marker

Gene Ruler™ DNA Ladder Mix
Fermentas, Germany

